

DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

BASIC PRINCIPLES OF ASSET PRICING THEORY: EVIDENCE FROM LARGE-SCALE EXPERIMENTAL FINANCIAL MARKETS

Peter Bossaerts

California Institute of Technology
and CEPR

Charles Plott

California Institute of Technology



SOCIAL SCIENCE WORKING PAPER 1070

February 2000

Basic Principles Of Asset Pricing Theory: Evidence From Large-Scale Experimental Financial Markets

Peter Bossaerts

Charles Plott

Abstract

Abstract: We report on six large-scale financial markets experiments that were designed to test two of the most basic propositions of modern asset pricing theory, namely, that the interaction between risk averse agents in a competitive market leads to equilibration, and that, in equilibrium, risk premia are solely determined by covariance with aggregate risk. We designed the experiments within the framework suggested by two theoretical models, namely, Arrow and Debreu's complete-markets model, and the Sharpe-Lintner-Mossin Capital Asset Pricing Model (CAPM). This framework enabled us to measure how far our markets were from equilibrium at any point in time, thereby allowing us to gauge the success of the models. The distance measures do not require knowledge of the (uncontrollable) level and dispersion of risk aversion among subjects, and adjust for the impact of progressive trading on the eventual equilibrium. Unlike in our earlier, thin-markets experiments, we discovered swift convergence towards equilibrium prices of Arrow and Debreu's model or the CAPM. This discovery is significant, because subjects always lacked the information to deliberately set asset prices using either model. Sometimes, however, the equilibrium was not found to be robust, with markets readily veering away, apparently as a result of deviations of subjective beliefs from objective probabilities. Still, we find evidence that this did not destroy the tendency for markets to equilibrate as predicted by the theory. In each experiment, we formally test and reject the hypothesis that prices are a random walk, in favor of stochastic convergence towards CAPM and Arrow Debreu equilibrium.

JEL classification numbers: G12, C92, D59

Key words: General Equilibrium Theory, Capital Asset Pricing Model, Complete Markets, Experimental Economics, Efficient Markets Hypothesis, Price Discovery

Basic Principles Of Asset Pricing Theory: Evidence From Large-Scale Experimental Financial Markets[¶]

Peter Bossaerts and Charles Plott

1 Motivation

Since the early fifties, two major considerations have driven the development of asset pricing theory, namely, (i) the presence of risk-averse investors, and (ii) asymmetric information. The former consideration has led to a better theoretical understanding of risk and its equilibrium compensation. The predictions from the theory are widely applied in practical areas, such as capital budgeting and portfolio performance evaluation, yet empirical support from econometric analysis of historical data has been dubious at best. The disconnect between theory and empirical support calls for an experimental investigation of some of the possible reasons.

Given the controversy, we decided to begin our experimental research with the simplest case: standard, symmetric-information asset pricing theory. This case already raises some difficult questions regarding experimental design and analysis. The presence of asymmetric information introduces a whole set of new issues into standard asset pricing theory. While the theoretical implications of asymmetric information are by now fairly well understood, empirical verification has proven to be difficult, and, hence, has been very limited.

The main premises of standard asset pricing theory have been that expected utility maximizing agents meet in a competitive marketplace, and that prices react instantaneously to equilibrate to a general equilibrium. The main implications have been that equilibrium prices will only reflect compensation for aggregate, i.e., nondiversifiable, risk, and that this risk is to be measured in terms of the covariance between the return on a security and some aggregate risk factor. Preferably, the latter can be identified in terms of the return on a specific (set of) benchmark portfolio.

The empirical support for the predicted relationships between risk as defined by the theory, and the ensuing equilibrium expected return, however, has not been convincing. Statistical regularities do appear in the cross-section

[¶]Financial support was provided by the Caltech Laboratory of Experimental Economics and Political Science, the National Science Foundation, and the International Center For Finance at Yale University. We thank Bill Brown (Claremont), Peter De Marzo (Berkeley), Will Goetzmann (Yale), Mark Johnson (Tulane), Tony Kwasnica (Penn State), Claude Montmarquette and Claudia Keser (CIRANO, Montréal), William Sharpe (Stanford), Woody Studenmund (Occidental College), and Ivo Welch (UCLA), for allowing us to involve their students in our experimental financial markets. The paper benefited from comments in a joint UCLA-Caltech workshop, at CEPR's 1999 Financial Markets Summer Symposium in Gerzensee, and during seminars at the Norwegian School of Management, the Claremont Graduate University, the University of Antwerp and the University of Louvain. E-mail: pbs@rioja.caltech.edu, cplott@hss.caltech.edu; Phone: (626) 395-4028; (626) 395-4209.

of average returns, but they seem not to be related to the theoretical predictions in any obvious way. See, e.g., Hansen and Singleton [1982], Fama and French [1992].

Despite an abundance of data, however, empirical tests require auxiliary assumptions that are not part of the general theory, but that are needed for the theory (including the auxiliary assumptions) to become falsifiable. Of course, rejections leave the researcher wondering whether they were caused by failure of some of the general economic principles, or by violations of some of the auxiliary assumptions.

Two examples of auxiliary assumptions are: (i) investors' ex ante beliefs coincide with the ex-post frequencies with which uncertain events (e.g., dividends) are realized, (ii) the pricing benchmark portfolio (the "market" in the case of the Capital Asset Pricing Model – CAPM) must be observable. The complexity of the economic environment surrounding historical financial markets makes us doubt that either assumption is and will always be met.

Experiments, however, may shed some light on the general validity of the basic principles of modern asset pricing theory, as well as the conditions for them to emerge, because experiments allow one to observe their workings in tightly controlled settings. Since the economic principles are supposed to be universal, they must predict the outcomes even in situations that are as simple as the ones in the experimental financial markets on which this paper will report. Failure for them to emerge ought to cast serious doubt on the general validity of the theory.

The ability to evoke the principles of modern asset pricing theory in the laboratory of course does not immediately imply that they can readily be used to predict outcomes in naturally occurring financial markets such as the NYSE or NASDAQ. The complexity of the environment facing the NYSE and NASDAQ, as well as the lack of observability of key parameters, inhibit attempts at extrapolation.¹

2 Issues Of Particular Concern In Experimental Tests Of Asset Pricing Theory

Experimentation has recently become an important means in determining the reliability of many aspects of economic theory, to the point that it has become an invaluable source of information in the design of markets. Virtually all experimentation, however, has been focusing on situations where the theoretical outcomes can easily be derived, albeit often at the expense of strong auxiliary assumptions such as risk neutrality. The advantage has been that prices can be computed explicitly, such that the success of the experiments can directly be measured in terms of the distances between observed and theoretical price levels. Examples of such experiments in the area of financial economics include Forsythe, Palfrey and Plott [1984], Plott and Sunder [1988], Smith, Suchanek and Williams [1988], Forsythe and

¹Analogously, an understanding of the biochemistry of isolated processes within cells does not allow one to readily predict the behavior of complex cellular structures such as human beings.

Lundholm [1990], Camerer and Weigelt [1993], Lei, Noussair and Plott [1998].

With respect to multi-security financial markets, however, there are no clear predictions one can make at the level of prices. Risk aversion is an integral part of the theory, and, in fact, asset pricing theory's *raison d' être* of financial markets, as mentioned before. Hence, simply assuming risk neutrality would invalidate the main premise on which the theory is built. Unfortunately, absent knowledge of the actual risk aversion among subjects in the experimental financial market, the theory does not allow one to compute prices. Worse even, there is in general a serious problem of multiplicity. In the case of the Capital Asset Pricing Model, the reader is referred to Nielsen [1988] for examples.

In such a situation, some experimentors would rather assume that subjects are risk-neutral, and *induce* a particular pattern of risk aversion for which unique theoretical price levels can be computed. Example: Forsythe, Palfrey and Plott [1984]. The presumption that subjects are risk-neutral is suspect, however, and, at a minimum, requires experimental verification. Indeed, in the experiments we are about to report, as well as in earlier experiments, there is clear evidence of (i) risk aversion among subjects, (ii) differences in risk aversion across subjects. This may in part have been caused by our deliberate usage of leverage in order to magnify the risks of the investments, as we will discuss later.

In fact, we observe a level of risk aversion in the experimental financial markets that, if extrapolated, would lead to extreme hesitance to take on any of the risks that subjects actually face in their everyday surroundings. Granted, the enormous historical risk premia on, e.g., U.S. common stock, can be interpreted as reflecting the extreme risk aversion that would be required to reconcile the outcomes in experimental and naturally occurring financial markets. See Mehra and Prescott [1985]. But we refrain from extrapolating our observations, because, as mentioned before, naturally occurring financial markets face a far more complex environment of which we only have a minimal understanding.²

We start from the premise that risk aversion may be present in experimental financial markets, and even be the driving factor behind the trading. But this new premise requires a new methodology. The success of the experiments cannot simply be measured anymore in terms of the distance of observed from predicted prices. So, other *measures* of the distance between actual and predicted outcomes needed to be developed. At the same time, an experimental setting had to be worked out in which these measures could most easily be computed.

The incorporation of risk aversion into the design and analysis of multiple financial markets experiments not only complicates the measurement of the distance of the market from equilibrium at any point during the experiment. It also introduces another difficulty which at first would seem insurmountable without appeal to ingenious experimental designs.

Unlike assumed in the theoretical development of asset pricing models, financial markets clearly do not equilibrate instantaneously (if they do at all – but this question is the subject of our investigation). We will refer to the process of equilibration as the *price discovery process*. The presence of an equilibration process is obvious from a casual inspection

²E.g., the environment in our experimental financial markets is clearly stationary, whereas that surrounding naturally occurring financial markets may very well be nonstationary.

of experimental data, even if it is ignored in the econometric analysis of historical price data from naturally occurring markets. Along the path of price discovery, however, trades take place, which change endowments, and, hence, desired future trades.

The out-of-equilibrium trading may shift endowments to the extent that the original equilibrium cannot be reached anymore. Hence, the equilibria that the experimenter envisaged on the basis of the initial allocations will not necessarily emerge in the actual experiments. The experimenter then runs the risk of rejecting the theoretical predictions, not because they are wrong, but because potential endowment shifts during the price discovery process were not accounted for in the formulation of the theoretical predictions.

Consequently, we need a clear measure of how far the market is from equilibrium at any point in time, that takes into account, (i) unknown patterns of risk aversion across subjects, (ii) unpredictable shifts in subjects' endowments along the price discovery path.

3 Novel Design And Analysis To Accomodate Asset Pricing Theory Applied To Experimental Markets

At the core of modern asset pricing theory is the hypothesis that financial markets *equilibrate* (instantaneously) and that equilibrium risk premia can be characterized in terms of the *covariation of an asset's return with aggregate risk*. This, of course, will not necessarily obtain in a situation of asymmetric information, so we deliberately excluded that possibility in the experimental design.

We can use mathematical language. Let R_i denote the return (payoff divided by price) on asset i and let $u'(C)$ denote the marginal utility of consumption of some "aggregate" investor (if it exists). We have a one-period model, so all wealth is consumed at the end, and, hence, "consumption" and (terminal) "wealth" can be used interchangeably. From the first-order conditions for optimality of the aggregate investor's investment-consumption plan, it must be that:

$$E[u'(C)R_i] = \lambda$$

(for some $\lambda > 0$). Provided there is a riskfree security with return R_F , this can be restated in terms of covariances, as follows:

$$E[R_i - R_F] = -\text{cov}(R_i, \frac{u'(C)}{E[u'(C)]}). \quad (1)$$

The task of asset pricing theory has been to identify the aggregate consumer (and to prove that it exists), so that $u'(C)$ can readily be measured. If measurement of aggregate consumption is an issue, aggregate wealth could be used instead. (Terminal) aggregate wealth equals the payoff on the aggregate portfolio of all investments in the economy, which has become known as the *market portfolio*.

It deserves emphasis that (1) generates a set of equations that define the Walrasian equilibrium in financial markets. That is, they produce a set of (nonlinear) equations, to be solved for prices. The latter are implicit in (1). To make them explicit, let X_i denote the payoff on asset i , let P_i denote its price, let P_F denote the price of the riskfree security, and assume that the latter's payoff is 1. Then:

$$E\left[\frac{X_i}{P_i} - \frac{1}{P_F}\right] = -\text{cov}\left(\frac{X_i}{P_i}, \frac{u'(C)}{E[u'(C)]}\right).$$

Because such equations are highly nonlinear (hyperbolic) in the unknown prices, existence and multiplicity problems should not come as a surprise.

In words, (1) predicts that equilibrium securities prices will be such that risk premia (expected excess returns) are determined by covariation with aggregate risk, where aggregate risk is to be measured in terms of the marginal utility of an aggregate investor. This implies not only that risk premia will exist, but also what their form will be. Let us formally state the first part, and then study more closely the second part.

- **Risk Pricing Property:** Risk premia will exist, i.e., equilibrium prices will generally deviate from expected payoffs.

There are at least two (mutually consistent) situations in which the prediction in (1) acquires more precise content. First, if *markets are complete* (in the sense that the number of independent securities equals the number of states of nature), Arrow-Debreu equilibrium theory predicts a specific pattern for the prices of Arrow-Debreu securities for states with equal likelihood, namely, that their ranking be inverse to the ranking of the aggregate consumption (wealth) in the economy. (Debreu [1959].)

To see how this follows from (1), let there be S possible states, indexed $s = 1, \dots, S$, each with probability π_s ($\sum_{s=1}^S \pi_s = 1$). Let P_s denote the price of state security s (that is, the security that pays \$1 if state s occurs, and \$0 otherwise). Let C_s denote the consumption (wealth) of the aggregate investor (who exists, by virtue of complete markets) in state s . Re-arranging (1), one obtains:

$$P_s = P_F \pi_s \frac{u'(C_s)}{E[u'(C_s)]}. \quad (2)$$

Therefore, two states s and s' with equal likelihood ($\pi_s = \pi_{s'}$) have prices that are ranked solely because of differences in consumption (wealth). If $C_s > C_{s'}$, then $u'(C_s) < u'(C_{s'})$ (because of decreasing marginal utility, i.e., risk aversion), and, hence, $P_s < P_{s'}$. The ranking of the state prices is the inverse of the ranking of aggregate consumption (wealth). This observation is the foundation for another theoretical property that we want to study in experimental markets.

- **State Price Property:** The ranking of state prices for equally likely states will be the inverse of the ranking of aggregate wealth in those states.

It is easy to design an experimental market with a complete set of securities markets. But the theoretical prediction is ordinal (involving mainly the rankings of Arrow-Debreu securities prices), and, hence, weak. So, the experimental setting needs to be specified further in order to generate a second, potentially cardinal, theoretical prediction.

The second situation where (1) acquires more precise content is one where the predictions of the Capital Asset Pricing Model (CAPM, developed in Sharpe [1964], Lintner [1965] and Mossin [1966]) can be invoked. This requires either normally distributed payoffs or quadratic utility.³ Quadratic utility is unlikely to provide a general description of subjects' global preferences. Still, quadratic utility may be a reasonable local approximation of subjects' actual preferences, given the limited stakes in our experiments.

Of course, one can never be sure *ex ante* that quadratic utility provides a reasonable local approximation of subjects' preferences. But neither can one assume that subjects are expected utility maximizers, or that markets equilibrate, both of which are key components of the theory that is to be studied experimentally. The actual experimentation will provide support, and, if not, allow the researchers to change the parameters, with the aim of determining the circumstances under which the theoretical predictions tend to emerge.

Consequently, our experiments were designed around the CAPM. This model predicts that risk premia will be proportional to the covariance between a security's return and the return on a specific benchmark portfolio, namely, the market portfolio (i.e., the value-weighted portfolio of all risky securities).

To see this mathematically, re-consider (1). Utility is quadratic, and, hence, $u'(C) = a - bC$, for two positive coefficients a and b . (Again, C can be interpreted interchangeably as aggregate consumption or aggregate wealth.) Re-arranging (1), one obtains:

$$E[R_i - R_F] = \frac{b}{E[u'(C)]} \text{cov}(R_i, C).$$

Let P_M denote the price of the market portfolio (the aggregate portfolio of investments in the economy). Its payoff, X_M , equals C . Let R_M denote its return, i.e., X_M/P_M . For the market portfolio,

$$E[R_M - R_F] = \frac{bP_M}{E[u'(C)]} \text{var}(R_M)$$

Using this to solve for $b/E[u'(C)]$, one obtains:

$$E[R_i - R_F] = \text{cov}(R_i, R_M) \frac{E[R_M - R_F]}{\text{var}(R_M)}. \quad (3)$$

In words, risk premia are proportional to covariances with the return on the market portfolio. The constant of proportionality is given by the expected excess return on the market portfolio divided by its variance.⁴

³One cannot implement normally distributed payoffs when at the same time requiring complete markets.

⁴Conventionally, the CAPM is re-written as follows:

$$E[R_i - R_F] = \beta_i E[R_M - R_F],$$

There is a simple way to measure whether this prediction is satisfied in the experiments, involving the *Sharpe ratio* (reward-to-risk) ratio of the market portfolio, namely,

$$\frac{E[R_M - R_F]}{\sqrt{\text{var}(R_M)}}.$$

All risk premia will be proportional to the covariances with the market return if and only if the Sharpe ratio of the market is maximal, i.e., equals the maximum possible reward-to-risk ratio in the marketplace. That is, the property in (3) will hold for all assets *if and only if* the market's Sharpe ratio is maximal. In finance jargon, this means that the market must be *mean-variance efficient* (generating minimum volatility given its mean return). For a precise statement of this result, see Roll [1977].

Away from the CAPM equilibrium, the market's Sharpe ratio will not be maximal. Hence, the difference between the market's Sharpe ratio and the maximal Sharpe ratio provides a simple measure of how far the market is from the CAPM equilibrium. We will exploit this feature to determine the success of our experiments. That is, we will investigate whether the following property obtains in our experiments.

- **Sharpe Ratio Property:** The Sharpe ratio of the market will be maximal.

The CAPM has one very important advantage. It provides a *generic description* of financial market equilibrium, irrespective of the distribution of risk aversion across investors. Moreover, this description is independent of investors' endowments. More specifically, the property that risk premia will be proportional to covariances with the market return is common to all CAPM equilibria and does not require one to know either the pattern of risk aversion or the distribution of wealth across subjects.

For this reason, the CAPM is an ideal framework within which to study experimental financial markets. It does require quadratic utility, but, again, experimentation should give the researcher a feeling for the parameter configuration that is needed for this to be a sensible local approximation to subjects' actual preferences.

As pointed out before, the empirical support for the CAPM in econometric analysis of historical data from naturally occurring markets is weak at best. But that does not disqualify the CAPM as a framework for the design and analysis of experimental financial markets. For one thing, the environment of naturally occurring financial markets is far more complex, and the experimental setting can be purposely built to give the CAPM its best chance to emerge. Moreover, the econometric analysis of historical data is marred with problems. We mentioned two problems earlier on, arising from the lack of observability of the market portfolio, and from the (debatable) necessity to assume that the market's ex-ante beliefs coincide with ex-post frequencies.

where

$$\beta_i = \frac{\text{cov}(R_i, R_M)}{\text{var}(R_M)},$$

i.e., the slope coefficient ("beta") in an OLS projection of R_i onto R_M .

We know of very few experiments that have been designed around the CAPM. Kroll, Levy and Rapoport [1988] mainly studied individual portfolio choice (in particular, whether individual holdings could be separated into a position in the riskfree security and a position in the market portfolio). Levy [1997] studied the CAPM in a repeated call market with twenty securities and discovered a significantly positive relationship between mean excess return and covariance with the market return, providing partial support for the CAPM. (As mentioned before, the CAPM actually imposes a much stronger restriction, namely, that this relationship be proportional.) Bossaerts, Kleiman and Plott [1998] discusses results of a set of experiments based on a design that is almost identical to that of the ones described here. It reports slow convergence towards the CAPM. In addition, the price discovery process generally stopped short of the equilibrium. The authors conjectured that thin trading slowed down and inhibited full convergence. The thick-market experiments described below were set up to test this conjecture.

4 On The Significance Of Experimental Tests

Some may wonder why we insist on designing the experiment with an explicit asset pricing model in mind. Would our attempts to give the CAPM its best chance to emerge not make its manifestation self-explanatory? Why not design the markets in an arbitrary way, thereby giving our tests maximum “power”?

Such reasoning, however, reflects a blind belief in one of the most fundamental concepts of neoclassical economics, namely, that markets equilibrate. There are at least two reasons to suspect that *financial markets in particular may not equilibrate*, even if this very statement sounds unorthodox.⁵

First, financial market equilibrium requires that there be no arbitrage opportunities (see Harrison and Kreps [1979]). While the converse is not necessarily true, arbitrage opportunities may arise along a market’s price discovery path. In earlier experiments, we indeed observed many arbitrage opportunities. The presence of such profit opportunities may keep investors from doing what equilibrium theory requires them to do, namely, to trade up to desirable portfolios. Instead, they may be distracted by arbitrage, thereby sending signals to the market which do not reflect their risk aversion and eagerness to re-allocate risk, but merely their eagerness to make money. This could eventually destroy the price discovery process itself.

If such arbitrage opportunities do arise, and the market ultimately does reach its equilibrium, one wonders what keeps it from being distracted. In earlier experiments, we observed, however, that many of the arbitrage opportunities do not represent “good deals,” in part because they are not scalable, unlike assumed in the theory.

There is a branch of unorthodox economics that rejects the hypothesis that markets equilibrate, namely Neo-Austrian economics. Whereas it attributes lack of equilibration to vastly different causes (nonstationarities *vs.* arbi-

⁵Experimental evidence that systems of multiple goods markets converge, albeit slowly, can be found in Noussair, Plott and Rezin [1995].

trage opportunities; see the discussion in Benink and Bossaerts [1998]), Neo-Austrian economics does share the concern that neo-classical theory overemphasizes optimal investment decisions in equilibrium, at the expense of analyzing what an investor ought to do out of equilibrium. This bias obviously originates in the presumption that financial markets equilibrate instantaneously.

Second, the stability of financial market equilibria must be questioned. This is certainly so in the CAPM, where risky securities are complementary at the equilibrium (the more one wants of one, the more one allocates to another, as a hedge). It is long known that equilibria with complementary goods create serious stability concerns, in the sense that simple tatonnement processes may inhibit convergence once a deviation from the equilibrium has taken place. Theoretically, markets may not converge back to any equilibrium at all. See Arrow and Hahn [1971] for a general discussion. For an experimental study of the phenomenon of nonconvergence, see Anderson, Granat, Plott and Shimomura [1999]

Of course, the price discovery process in our experimental financial markets is far from tatonnement, because of endowment effects whenever there is a trade. The experiments should reveal whether the speculative opportunities that cycling generates suffice to force the market to eventually converge to an equilibrium.

The significance of our experimental results can be appreciated differently, though. Even if subjects had known the purpose of our experiments (to study predictions from general equilibrium asset pricing theory), and had liked to “help” us by all purposely investing according to the theory, thereby generating the results we had hoped for, they would not have been able to do so. For instance, one way to ensure that the CAPM emerges is for subjects to all buy combinations of the riskfree security and the market portfolio. This requires knowledge of the composition of the market portfolio, however, which subjects were not told, and, as comments from subjects after the experiment revealed, which could not be extracted from the data. Likewise, to price Arrow-Debreu securities according to the theory (securities for equally likely states with lower aggregate endowment carrying a higher price), subjects had to know the aggregate endowment across states, which they were not told.

Therefore, the experimental confirmation of the predictions of modern asset pricing theory is neither a foregone logical conclusion nor a self-fulfilling prophesy.

5 Large *vs.* Small-Scale Experimental Environments

This paper reports on six large-scale, web-based financial markets experiments at Caltech, involving MBA students at Berkeley, Stanford, Tulane, UCLA and Yale, as well as undergraduate and graduate students from Caltech, Claremont, Montréal and Occidental. We had previously run similar experiments based on the time-tested MUDA framework. Reports on these experiments can be found in Bossaerts, Kleiman and Plott [1998]. Basically, while these experiments did reveal convergence towards CAPM equilibrium, the process was slow and apparently almost invariably aborted

short of the equilibrium. After these experiments, we conjectured that thin markets were at the root of the convergence problems, because they hinder subjects' attempts to re-allocate their portfolios appropriately in an environment where trades in one market cannot be conditioned on events in other, parallel markets.

Since the earlier MUDA-based experiments, two routes have been attempted to investigate the thin-markets problem. First, we have been (and still are) investigating new market designs that explicitly facilitate portfolio trading. Second, we replicated the earlier MUDA experiments in far thicker markets, run on the internet with Caltech's Marketscape software. It are these experiments that the present paper reports on.

Our conjecture that the thickness of the market plays a role in equilibration was supported in the large-scale experiments. As will be documented later on, the increase in the number of subjects led to a significant reduction in the average bid-ask spread and its volatility, making it far easier for subjects to rebalance portfolios, and, consequently, facilitating equilibration.

6 Experimental Parameters and Procedures

All six experiments were organized as follows. Three assets could be traded. Two, labeled Securities A and B, had a risky payoff, determined by the random drawing of one of three "states," referred to as States X, Y and Z. The third asset, labeled Note, was riskfree, and, unlike the risky securities, could be sold short, up to eight units. (In the CAPM equilibrium, shortsale of the risky securities has no bite, whereas the possibility to shortsell the riskfree security is potentially important.) Subjects were endowed with a certain number of Securities A and B. The Notes were in zero net supply.

Subjects could offer for sale or purchase, or trade, in each of the securities during periods of fifteen to twenty-five minutes, after which the state was announced and dividends were paid depending on the final holdings of the securities. The assets were then taken away, and subjects were given a fresh supply of the risky securities, as well as a certain amount of cash. The time between initial allocation of securities and cash and final realization of the dividends will be referred to as *periods*. Each experiment involved up to eight periods.

Cash was given at the beginning of every period. At the end of each period, subjects had to pay the experimenter from the dividend proceeds and the final cash for the initial cash as well as the initial allocation of risky securities. This payment was referred to as "loan repayment." This loan repayment creates leverage, causing a magnification of the risk involved in the holding of Securities A and B.

All accounting was done in terms of a fictitious currency called francs, to be exchanged for dollars at the end of the experiment at a pre-announced exchange rate. In the last three experiments, subjects were also given an initial sign-up fee, theirs to keep if they successfully finished the experiment.

Subjects were barred from further trading if they ran net cumulative losses more than two periods in a row

(although we generally gave subjects a third chance).

The payoff matrix for all experiments remained the same, namely:

State	X	Y	Z
Security A	170	370	150
Security B	160	190	250
Notes	100	100	100

The remaining data and parameters for the experiments are displayed in Table 1. Notice the variation in initial endowments across experiments. In the CAPM equilibrium, subjects should hold combinations of the market portfolio and the riskfree security, depending on their risk aversion. So, in principle, there would be more trade when subjects start off with a portfolio of risky securities that is substantially different from the market portfolio. On the other hand, the trading that is required for subjects to re-allocate their holdings towards the market portfolio may confuse the market, and the CAPM may not emerge. To understand this trade-off, we varied the initial allocations.

There is another reason why we varied the allocations. We wanted to avoid that (Caltech) subjects who participated more than once would have learned the composition of the market portfolio and be able to price the securities according to the CAPM (if they really wanted to do so). The purpose of the experiments is to induce the CAPM through economic forces, and not through deliberate pricing on the basis of a theoretical model.

When subjects were given different endowments, the parameters (including the loan repayments) were set such that subjects had similar risk/return tradeoff if they traded up to mean-variance optimal portfolios at one of the possible equilibrium price sets (i.e., a set of prices such that the market portfolio had the highest possible Sharpe ratio). Also, the loan repayments were set such that a subject wishing to unload all risky securities would make some money, even if minimal. Likewise, a subject wishing to invest all wealth in Security A (which generally constituted a mean-variance optimal, albeit risky, investment) would approximately double income relative to a 100% investment in the market portfolio.

The actual payoffs obviously depended on the sequence of states that happened to be drawn during the experiment. Because each experiment involved only up to 8 periods, the frequency of occurrence of each state deviated from the ex ante probabilities ($1/3$ for each state), but all states ended up occurring in each experiment. While subjects were told that states were random and equally likely, some of them clearly ignored this and started gambling on the occurrence of a state that had not appeared yet. This apparently affected prices in two experiments (referred to as the 990407 and the 991111 experiments).

We had little control over the number of participants in the experiments. Subjects had to register in a database several days before the actual experiment. After that, they could retrieve an ID and password with which to log on to the Marketscape website from which the experiment would be run, and execute some trades as practice. In general,

however, subjects were slow to register. Some even decided to forego the opportunity to practice before the actual experiment and registered after the practice session was closed.

Hence, the number of subjects was random. In one experiment, we only had 19 subjects (the one referred to as 990211). It is interesting to study the effects of the consequent thin trading.

Subjects were given clear instructions, which they could read on the website for each experiment, and which we did not alter across experiments. These instructions included a description of some portfolio strategies. They did not mention a strategy whereby the subject is invited to gamble against the experimenter's random drawing of states.

It should be pointed out that all subjects had familiarity with financial markets. This is obvious for the MBA students. The Caltech students had been exposed to financial markets through a class in introductory finance, often complemented with a class in investments analysis. Claremont and Occidental students were taking economics/econometrics classes.

The websites for the six experiments are as follows:

<http://eeps2.caltech.edu/market-981007yale> ("981007")
<http://eeps2.caltech.edu/market-981116uclacit> ("981116")
<http://eeps2.caltech.edu/market-990211> ("990211")
<http://eeps2.caltech.edu/market-990407> ("990407")
<http://eeps2.caltech.edu/market-991110> ("991110")
<http://eeps2.caltech.edu/market-991111> ("991111")

The interested reader can visit these websites, read the instructions, inspect the trading interface, and display the trading history (including pricing).⁶

7 Results

Asset pricing theory states that financial markets exist to re-allocate risk among risk averse investors. Moreover, financial markets equilibrate to a point where risk premia can be characterized in a specific way: they are determined solely by covariance with aggregate risk [see (1)].

As discussed in the theory section, the presence of risk aversion generates specific predictions, which we repeat here.

- **Risk Pricing Property:** Risk premia will exist, i.e., prices will generally deviate from expected payoffs.
- **State Price Property:** The ranking of state prices for equally likely states will be the inverse of the ranking of aggregate wealth in those states.

⁶Anonymous login requires the ID 1 and password a.

- **Sharpe Ratio Property:** The Sharpe ratio of the market portfolio will be maximal.

Before we proceed to present experimental results, it should be pointed out that there is ambiguity about the meaning of “equal likelihood” as well as the probabilities with which the Sharpe ratios have to be computed. In theory, it suffices that these are identified with investors’ (supposedly common) *subjective beliefs*. There may be a discrepancy between investors’ subjective beliefs (ex-ante expectations) and the objective probabilities (ex-post frequencies) with which uncertain events factually occur.

We commented earlier in the paper that this may be one of the reasons why empirical tests of asset pricing theory on historical field data have not been supportive in general. Such tests invariably assume that subjective and objective beliefs coincide. The empiricist’s task is limited to handling error in the estimation of the objective probabilities.

In our experimental setup, objective probabilities were controlled. For simplicity, we made the states equally likely, and random over time, and we announced this to the subjects. Even so, subjective beliefs may still differ from the objectively announced probabilities, in part because subjects do not understand randomness⁷ or because they truly believe there is not such a thing as a random sequence.

Consequently, experimental observation may deviate from theoretical predictions because we implicitly impose that subjective beliefs equal objective probabilities. To the extent that we did observe anomalies, we investigated if these could be explained in terms of deviations between subjective and objective probabilities.

To the above predictions, we add a conjecture, which is not part of standard asset pricing theory (because it focuses on equilibrium, and not equilibration), but which was suggested as a conjecture after analysis of earlier, thin-markets experiments.

- **Conjectured Bid-Ask Spread Property:** Financial market equilibration will be facilitated by market thickness, because an increase in the number of subjects will lead to a reduction in the size of the bid-ask spread.

Since the number of participants in our experiments varied substantially, we are in a position to explore our conjectured property.

The evidence on the above properties will be presented, temporarily ignoring issues about beliefs. We will first discuss the dynamics of securities prices, which will enable us to establish Risk Pricing Property in the data. Subsequently, we discuss the evolution of the Sharpe ratio of the market portfolio, in order to measure to what extent our markets tended towards Sharpe Ratio Property. Next, we present the dynamics of state security prices, to verify whether our markets were characterized by State Price Property. Finally, we discuss the effect of market thickness (number of subjects) on the bid-ask spread and market equilibration (Conjectured Bid-Ask Spread Property).

⁷Randomness is sometimes confused with equal frequency in finite samples.

7.1 Price Dynamics: Evidence For Risk Pricing Property

Figure 1 displays the evolution of the prices of the three securities in the first experiment, 981007. Each observation corresponds to a trade in one of the three securities (the price of the other two securities was set equal to their last trading price). The picture is representative for what happened in all the experiments, not just in 981007.

The first observation one can make is that prices almost always deviate from expected payoffs (230, 200 and 100, for Securities A, B and Notes, respectively). This suggests support for Risk Pricing Property, and reveals that risk aversion plays an important role in experimental financial markets, and, hence, must be taken into account explicitly when assessing the success of the experiments.

There is a striking cyclicity in the prices of the Notes. Prices start out low in each period, but converge to 100 (the payoff), implying that the riskfree rate decreases steadily over the course of a period. Because there is no time value of money, the equilibrium interest rate must be zero. One generally observes near zero interest rates at the end of each period. The transient positive interest rates reflect subjects' willingness to borrow in order to execute the securities trades that are needed to optimize their portfolio. Because purchases are paid for in cash, and orders in one market cannot be made conditional on trades in other markets, subjects need to have a certain amount of cash to execute their purchase orders if they do not wish to sit there and wait for their sales orders for other securities to be executed first.

As a matter of fact, monetary theory has appealed to such a *cash in advance* constraints to explain the presence of interest-bearing nominally riskfree securities along with zero-interest cash. (Clower [1967], Lucas [1982].) Monetary theorists, however, have generally attributed cash-in-advance constraints to the purchase of consumption goods. In our experiments, they are clearly at work in investment decisions as well. Our findings would suggest that it may be worthwhile to study the effect of portfolio re-allocations in explaining the behavior of interest rates in naturally occurring financial markets as well.

A final observation concerns the level at which the prices of Securities A and B settle at the end of each period. It appears to differ across periods and experiments. Do this represent multiple equilibria? As mentioned before, we can answer this questions in two ways. Because the experimental markets were purposely designed to be complete (the number of securities equals the number of states), we can imply the prices of Arrow-Debreu securities from the prices of the traded securities and verify whether their ranking is as theory predicts. Second, we can compute the Sharpe ratio (reward-to-risk ratio) of the market portfolio at any point in time, and compare it to the maximum reward-to-risk ratio that one could obtain in the marketplace. Because the latter directly measures the distance of the market from (CAPM) equilibrium, we present it first.

7.2 Market Sharpe Ratios: Evidence For Sharpe Ratio Property

Figures 2, 3, 4, 5, 6 and 7 display the evolution of the difference between the Sharpe ratio (mean return divided by standard deviation) of the market portfolio against the maximum possible Sharpe ratio.⁸ Each observation corresponds to a trade in one of the three securities (the price of the other two securities was set equal to their last trading price). In the CAPM equilibrium, the distance between the market and maximal Sharpe ratio is zero. Across the six experiments, there is a pronounced tendency for the markets to move toward the CAPM equilibrium. Almost invariably, when the markets are kicked off the equilibrium, they revert back towards it. Consequently, the experimental markets tend towards Sharpe Ratio Property.

We would like to emphasize again that there is an important sense in which this support for the CAPM is remarkable. Subjects are never told what the market portfolio is, yet, when we discover the CAPM equilibrium, their combined holdings of risky securities (i.e., the market portfolio) add up to a mean-variance optimal portfolio. Even if the subjects believe that they must invest according to the CAPM (perhaps because they were told in investments classes, and knew that we were testing for the CAPM), and want to buy the market portfolio, they would not be able to do so, because they do not know the composition of this portfolio. In other words, the emergence of the CAPM is entirely the result of financial market equilibration in the presence of risk averse investors, and not of subjects' deliberate attempts to generate the results that support prevailing theory (the CAPM).

Collectively, the figures provide strong support for two fundamental principles of modern asset pricing theory, namely, that financial markets equilibrate, and that, in equilibrium, risk premia are determined by covariance with aggregate risk (which we measure as the return on the market portfolio). In contrast with the theory, however, the equilibration is far from instantaneous. Moreover, there is evidence that markets can easily be kicked off an equilibrium. This is most striking in the 990407 and 991111 experiments (Figures 5 and 7). These instabilities are bothersome. It is possible, however, to attribute this to subjects' speculation about the state, even if they had no information. This will be clear when we study the Arrow-Debreu securities prices. It is not obvious why such speculation was pronounced in the 990407 and 991111 experiments, and less so in other ones.

One can translate the visual evidence of Figures 2 to 7 into formal statistical statements. In particular, it is possible to reject the null hypothesis that price dynamics were a random walk against the alternative that price movements were determined by economic forces that pulled the markets towards CAPM equilibrium. The random walk hypothesis is an appropriate benchmark, because it is devoid of economic content beyond the statement that it does not allow for arbitrage opportunities. It has been the basis of much empirical research in finance, where it is known as the Efficient Markets Hypothesis. Under the random walk hypothesis, price changes are independent of past history, and, hence, speculation cannot be profitable. We would like to know whether there is more to financial markets than just the

⁸In computing the maximum Sharpe ratio, shortsale constraints on Securities A and B were accounted for.

inability to “beat” them. In particular, do they equilibrate, to the point that only aggregate risk is priced?

Even under the random walk hypothesis, the market can accidentally move towards the CAPM equilibrium. Under the alternative hypothesis, however, this movement is not accidental: the farther away from equilibrium, the more the market is attracted towards equilibrium. We wish to formally distinguish the two hypothesis, and to test whether the visual evidence in favor of the CAPM in Figures 2 to 7 is not a case of mere luck.

We take the random walk hypothesis as the null, and test it against the hypothesis that the market is pulled towards the CAPM. Our test works as follows. Let Δ_t denote the distance between the Sharpe ratio of the market and the maximum Sharpe ratio at transaction t . Δ_t is the absolute value of the variable plotted in Figures 2 to 7. Consider the projection of the change in Δ_t onto Δ_{t-1} :

$$\Delta_t - \Delta_{t-1} = \kappa \Delta_{t-1} + \epsilon_t. \quad (4)$$

where κ is such that ϵ_t is uncorrelated with Δ_{t-1} . CAPM implies $\Delta_t = 0$; convergence to CAPM pricing implies $\kappa < 0$. We then determine the distribution of the least squares estimates of κ under the null hypothesis of a random walk, by randomly drawing from (bootstrapping) the empirical joint distribution of changes in transaction prices. The null hypothesis of a random walk is rejected (in favor of stochastic convergence to CAPM) if the least squares estimate of κ is beyond a critical value in the left tail of the ensuing distribution. This testing procedure is a variation of *indirect inference* (see Gouriéroux, Monfort and Renault (1993)): we summarize the data in terms of a simple statistical model (in our case, a least squares projection) and determine the distribution of the estimates by simulating the variables entering the statistical model. Instead of simulating off a theoretical distribution, we bootstrap the empirical distribution, however.⁹

For each experiment, we estimated κ using OLS. 5% and 10% critical values under the random walk null hypothesis were determined by bootstrapping from the empirical joint distribution of price changes (we generated 200 price series of the same length as the sample used to estimate κ).¹⁰

Table 2 reports the results. The null of a random walk is rejected in each experiment. In the case of 991110, however, the rejection is marginal (above the 10% level). The latter may appear surprising, because the visual evidence from Figure 6 would make one believe that there is a strong case for the CAPM. Yet, the statistical analysis indicates that there is a high probability to accidentally obtain a plot like the one in Figure 6, even if prices were a random walk. The contrast between Table 2 and Figure 6 underscores the need for an appropriate, formal statistical test of the theory, even in an experimental setting.

The rejections of the random walk hypothesis reported in Table 2 do not imply that the subjects ignored profit opportunities from speculating on price changes. This is because our rejection of the random walk is based on

⁹Gallant and Tauchen (1996) also use indirect inference, but, instead of matching an arbitrary statistical model, they match the scores of the likelihood function.

¹⁰We bootstrapped the mean-corrected empirical distribution, in order to stay with the null hypothesis of a random walk.

information that subjects could not condition on, namely, the Sharpe ratio of the market portfolio. This ratio is not readily determined from the history of price movements. Computation requires knowledge of the composition of the market portfolio. As emphasized before, subjects did not know this, and, hence, could not determine in which way prices would move when the market was still out of equilibrium, even if they believed in the CAPM.

7.3 The Dynamics Of State-Price Probabilities: Evidence For State Price Property

As mentioned before, equilibrium implies that the ranking of Arrow-Debreu securities prices of equally likely states must be inverse to the ranking of the aggregate payouts. Let us study this now.

We deduced the Arrow-Debreu securities prices for the three states from the transaction prices of the three securities. As has become standard in mathematical finance and derivatives analysis, we normalized the prices to add up to one. It is well known that the resulting numbers are probabilities. Hence, one refers to them as *state-price probabilities*.¹¹

In fact, it is common nowadays in mathematical finance and derivatives analysis to imply these state-price probabilities from the prices of derivative securities or fixed-income securities traded in naturally occurring markets. They can readily be used to price other derivative securities, or fixed-income securities with different maturities. It suffices that there be no arbitrage opportunities for the exercise to make sense (i.e., for the resulting numbers to be probabilities). See, e.g., Rubinstein [1994].

We will be the first, however, to be able to document that these numbers are not mere probabilities which would obtain under absence of arbitrage opportunities, but that they also have precise economic content, because they are directly derived from Arrow-Debreu securities prices. In equilibrium, the latter are constrained, and we will discover the constraints in our experimental financial markets.

Figures 8, 9, 10, 11, 12 and 13 display the evolution of the three state-price probabilities (one for each state). They were implied from the transaction prices of the traded securities. Each observation corresponds to a trade in one of the three securities (the price of the other two securities was set equal to their last trading price). Because states are equally likely, the equilibrium ranking of the state-price probabilities must be the inverse of the ranking of the aggregate payoff (payoff on the market portfolio). From the payoff matrix and the allocations in Table 1, it follows that state X must be most expensive, followed by states Z and Y, in all the experiments. So, the state-price probability for state Y must be lowest. etc.

One observes already early on in the 981007, 981116 and 991110 experiments (Figures 8, 9 and 12) how the state-price probabilities attempt to separate as predicted by the theory. The probability for State X generally increases to eventually become the highest, whereas that of State Y decreases to become the lowest, exactly as predicted by Arrow-Debreu general equilibrium theory.

¹¹Some refer to them as defining the equivalent martingale measure, or risk-neutral probabilities; see Harrison and Kreps [1979].

That State Price Property might appear in the data is far from obvious. Even if subjects had collectively wanted to price the state securities in accordance with the theoretical prediction, they would still have to know the aggregate endowment in each state, which they were not told, and apparently were unable to discover from the data.

The 990407 and 991111 results are most interesting. Figure 11 demonstrates how the 990407 state price probability of state Y is generally lowest, but that the ranking of states X and Z is often wrong, except near the end of periods 1, 2, 3, 4, 6 and 7 (the outcome in period 6 is not clearcut). In the 991111 case (Figure 13), the state price probability for X is generally highest, as predicted by the theory, but the ranking of the probabilities for Y and Z is reversed in periods 4, 5 and 6. We will come back to these state-price probabilities in a later section.

With the exception of a few periods in 990407 and 991111, the evidence suggests that state price probabilities moved in the direction predicted by Arrow-Debreu equilibrium. One would like a formal test to confirm this finding. In particular, one would like to determine whether indeed state price probabilities adjust towards Arrow-Debreu equilibrium when their ranking is not as predicted by the theory. As in the previous section, we take random walk pricing as our null hypothesis. That is, we determine what the probability is of observing the dynamics in state price probabilities in our experiments if prices were merely random walks (unpredictable, and, hence, unaffected by economic forces beyond absence of speculative profit opportunities). Under the alternative that markets are attracted by Arrow-Debreu equilibrium, we expect specific changes in the state price probabilities when they are not aligned appropriately. In particular, we expect the following.

Ranking of State Price Probabilities at t	Expected Effect
$P_{X,t} > P_{Y,t} > P_{Z,t}$	$(P_{Z,t+1} - P_{Y,t+1}) - (P_{Z,t} - P_{Y,t}) > 0$
$P_{X,t} > P_{Z,t} > P_{Y,t}$	Anything is Possible
$P_{Y,t} > P_{X,t} > P_{Z,t}$	$(P_{X,t+1} - P_{Y,t+1}) - (P_{X,t} - P_{Y,t}) > 0$
$P_{Y,t} > P_{Z,t} > P_{X,t}$	$(P_{X,t+1} - P_{Y,t+1}) - (P_{X,t} - P_{Y,t}) > 0$ or $(P_{Z,t+1} - P_{Y,t+1}) - (P_{Z,t} - P_{Y,t}) > 0$
$P_{Z,t} > P_{Y,t} > P_{X,t}$	$(P_{X,t+1} - P_{Y,t+1}) - (P_{X,t} - P_{Y,t}) > 0$
$P_{Z,t} > P_{X,t} > P_{Y,t}$	$(P_{X,t+1} - P_{Z,t+1}) - (P_{X,t} - P_{Z,t}) > 0$

These predictions are weak, because they only concern the sign of the change in the difference between two state price probabilities. The question is: are economic forces strong enough that the expected effects can be detected sharply?

As test statistic, we compute the frequency of observing (transiting to) the expected outcome for each state (ranking of state price probabilities). We subsequently average across states. The latter is mandated by the fact that, in finite samples, not all states need occur, in which case some transition frequencies are undefined. Let π denote the mean transition frequency. Notice that the second frequency will always be 1. We include this frequency, so that outcomes where the Arrow Debreu prediction holds ($P_{X,t} > P_{Z,t} > P_{Y,t}$) receive more weight. In analogy with our formal test

of Sharpe Ratio Property, we compute the distribution of π under the null hypothesis by bootstrapping the empirical joint distribution of price changes in each experiment (we generated 200 price series of the same length as the sample used to estimate π).¹²

Table 3 reports the results. In addition to the estimated mean transition frequencies (second column), we report 5%, 90% and 95% critical values under the null of random walk pricing. The striking aspect of the estimated mean transition frequencies is not only that they are highly significant, but also that they are remarkably constant across experiments. There are pronounced differences across experiments in terms of beginning prices and empirical distribution of price changes, which translate into marked differences in the distribution of π under the null of a random walk.¹³ Across all experiments, the estimated values for π are almost always the same, however, suggesting that the same forces are at work. They are in the right tail of the distribution under the null, and, except for 991110, significant at the 5% level. The weaker showing of 991110 confirms the finding in the previous section (see Table 2). Overall, Table 3 provides formal evidence that Arrow Debreu equilibrium predicts price movements in experimental financial markets better than the random walk hypothesis.

Again, the rejections of the random walk hypothesis reported in Table 3 do not imply that the subjects ignored profit opportunities from speculating on price changes. This is because our rejection of the random walk is based on information that subjects could not condition on, namely, the distribution of aggregate wealth across states.

7.4 Market Thickness, The Bid-Ask Spread And Equilibration: Evidence For Conjectured Bid-Ask Spread Property

From both the Sharpe ratios and the state-price probabilities, one can deduce that equilibration is hampered when markets lack the necessary thickness. We had observed equilibration problems in the earlier, thin-markets, MUDA experiments we ran (Bossaerts, Kleiman and Plott [1998]). In the 990211 Marketscape experiment, only 19 subjects showed up. While higher than in the MUDA experiments, this number was clearly insufficient.

In support of Conjectured Bid-Ask Spread Property, that market thickness led to lower bid ask spreads, and, hence, facilitates equilibration, Table 4 reports the average, median and standard deviation of the bid ask spreads in the four experiments, along with the number of subjects. The first period was always excluded in the computation of the statistics.¹⁴ Table 4 reports that the mean and median bid-ask spreads are uniformly highest in the 990211

¹²We bootstrapped the mean-corrected empirical distribution, in order to stay with the null hypothesis of a random walk.

¹³If the initial price configuration satisfies or is close to satisfying the Arrow Debreu equilibrium restriction, the simulated π s will be high (in the table above, the predicted outcome if $P_{X,t} > P_{Z,t} > P_{Y,t}$ obtains with unique frequency). This explains the high level of the 5% critical value for some experiments (990407 and 991110). This way, our procedure penalizes experiments that happen to start out with prices that (closely or exactly) satisfy Arrow Debreu equilibrium restrictions.

¹⁴Bid-ask spreads were computed only if there were both a standing bid and an ask, or when the bid or ask price was matched or improved upon; the bid-ask spread was set to be missing after a transaction (otherwise, an artificial increase would have been reported).

experiment, where only 19 subjects participated. With the exception of security B in the 991110 experiment, the volatility of the bid-ask spreads is also highest in 990211.

The results (bid-ask spreads, equilibration) of the 990211 experiment are qualitatively like in the earlier, thin-markets, MUDA experiments (Bossaerts, Kleiman and Plott [1998]). That is, they replicate our observations from equally thin markets, albeit with a different trading mechanism. The five other experiments employed a substantially larger number of subjects, and, altogether produce different results: overall smaller bid-ask spreads, as well as equilibration. The 991110 experiment is an outlier: it featured the highest number of subjects, yet the bid-ask spreads are not uniformly the lowest in least volatile. This illustrates that sampling error is definitely present, and, hence, that it would be dangerous to base inference on a single experiment.

Market thickness reduces the bid-ask spread. In thin markets, it is far more time consuming and costly to re-balance one's portfolio when the bid-ask spread is high. Apparently, this also inhibits market equilibration. See Figures 4 and 10. Hence, Table 4, together with the evidence from our earlier MUDA-based experiments, support our Conjectured Bid-Ask Spread Property.

7.5 Subjective Equilibrium

Let us turn to the state price probabilities in the 990407 and 991111 experiments.

First the 990407 results. Remember that Figure 11 demonstrated how the probability of state Y is generally lowest, but that the ranking of states X and Z was often wrong, except when the markets touch equilibrium in periods 1, 2, 3, 4, 6 and 7 (the outcome in period 6 was not clearcut).

The reversal of the ordering of the probabilities for states X and Z could be understood once we reveal the sequencing of the states. In the 990407 experiment, this sequencing was: X (period 1), Y, Y, X, Z, Y, X, Y (period 8). By the fourth period, some subjects had become convinced that period Z had to be next, because it had not yet been drawn (of course, this is incorrect), and, hence, they implicitly bid up the price of state Z. When state X actually occurred, they apparently became even more convinced that X would not occur anymore, and that Z had to be next (they happened to be right). By period 8, a similar speculation occurred. Since Z had only occurred once, some subjects were evidently convinced it had to be drawn next, interpreting incorrectly that "equal likelihood" means "equal frequency." Again, they implicitly bid up the price of state Z, making it hard for the market to move to a CAPM equilibrium.

One can suspect that long repetitions of our experimental setup will eventually teach subjects a more conventional meaning of randomness. (It is interesting to note that they made no such errors in the first and second periods.) One could argue that such subjects would not survive very long in naturally occurring financial markets,¹⁵ and, hence,

¹⁵As a matter of fact, four subjects were bankrupt and barred from further trading in the 990407 experiment because of repeated cumulative losses.

that the phenomenon we observe in our experimental markets would not be of practical importance. Unfortunately, the environment surrounding naturally occurring financial markets is sufficiently nonstationary (new inventions, new securities, new industries, new government ideologies, etc.) that one ought not to be easily convinced by this argument.

It appears that small deviations between subjective beliefs and objective probabilities can have dramatic effects on securities prices. In the context of the CAPM, Mehra and Sah [1998] have shown that slight changes in subjective preferences can indeed lead to large price movements. Our experiments provide indirect support for their theoretical calculations.

Asset pricing theory does not require that subjective beliefs be correct. So, its predictions about equilibration and the nature of equilibrium risk premia remain valid even in a world where subjective beliefs deviate from objective probabilities. The empirical methodology that we have used so far does require that subjective beliefs be identified with the objective probabilities. Despite our efforts to ensure that subjects' beliefs coincided with the objectively stated and constant probabilities, the price dynamics in our experimental markets appear to have been influenced in part by changes in subjects' expectations. But the goal of the experiments was not to verify whether subjects understood probabilities. Rather, we wanted to determine whether experimental financial markets tend to do what asset pricing theory predicts, namely, to equilibrate, to a point that risk premia are determined solely by covariation with aggregate risk.

In order to assess whether our markets continued to behave in accordance with asset pricing theory even after subjects' expectations drifted away from objective probabilities, we need a new methodology, to measure how far a market is from equilibrium, irrespective of its beliefs. A complete formulation is a formidable task, beyond the goals of this paper.

Still, we can derive a simple measure that would provide some indicative evidence of equilibration under subjective beliefs, as follows. The procedure relies on an explicit assumption about the beliefs reflected in market prices. Specifically, let the market conjecture (incorrectly) that states are drawn, *without replacement*, from an urn with four balls of type X, four of type Y, and four of type Z. This is meant to formalize the finding that subjects appeared to gamble over the nature of the next state to be drawn, as which we discussed above. The question is whether the market can be shown to equilibrate under such a hypergeometric, subjective belief structure when it could not under the objective probabilities.

Under the hypergeometric belief structure, states that occurred less frequently will be assigned higher chances of occurring next. For instance, state Z would have 50% chance after the fourth round, whereas states X and Y would occur only with 25% probability, having both been drawn twice before.

We recomputed the maximum Sharpe ratio and the Sharpe ratio of the market accordingly, and plotted the evolution of the difference. The bottom panel of Figure 14 displays this plot. For ease of comparison, the evolution of the Sharpe ratio difference based on objective probabilities is redrawn in the top panel (cf. Figure 5).

With the hypergeometric belief structure, the market evidently did equilibrate in all periods up to and including period 6, as is evident from the evolution of the difference between the (recomputed) Sharpe ratios of the market and the maximal one. See Figure 14.

The same analysis could be applied to the 991111 results. Remember that Figure 13 demonstrated how the probability of state X is generally lowest, but that the ranking of states Y and Z was often wrong, especially in periods 3, 4, 5 and 6. Again this may be explained in terms of speculation. The sequence of states in the 991111 experiment happened to be: Y (period 1), Z, Z, X, X, Y, Z, Z (period 8). From period 4 on, subjects speculated that Z was not going to be drawn soon anymore, and, instead, that Y, and, to a lesser extent, X, were to occur. When Y was not drawn, they became more insistent, until Y was eventually reached in period 6. The evidence from the state price probabilities is less clear, though, than in the 990407 experiment.

Instead, support that this explains the anomalies in the 991111 experiment can be obtained from the Sharpe ratios, once these are recomputed on the basis of a subjective belief model, whereby the market assumed that states were drawn from a urn with 12 balls (4 are labeled “X”, 4 are “Y” and 4 are “Z”), without replacement. Notice that this is the same subjective belief model that we already used to explain the 990407 anomalies. The bottom panel of Figure 15 displays the evolution of the difference between the subjective Sharpe ratio of the market portfolio and the maximum subjective Sharpe ratio. For comparison, the top panel replicates the evolution of the objective Sharpe ratios from Figure 7. Notice how the switch to subjective beliefs again demonstrates that markets did equilibrate even in periods 3 to 6.

8 The Pricing *vs* Allocation Paradox

The graphical evidence on Sharpe ratios and state price probabilities, and, perhaps even more convincingly, the formal statistical tests involving Sharpe ratios and state price probabilities, demonstrate that the CAPM and Arrow Debreu equilibrium successfully describe the dynamics of pricing in multisecurity experimental financial markets.

Both the CAPM and Arrow Debreu equilibrium make predictions about allocations as well. In field data, these are usually not verified, perhaps because of difficulties in measuring investors’ portfolios. Casual observation, however, has suggested that investors have historically not been holding the well-diversified portfolios prescribed by the CAPM. Neither have their investments led to consumption with perfect rank correlation across states, as predicted by Arrow Debreu equilibrium. In the laboratory, we have the opportunity to accurately observe portfolio choice, and rule out measurement error.

Let us focus on the CAPM and investigate to what extent this model describes actual portfolio holdings in experimental financial markets. In the CAPM equilibrium, investor would hold all risky securities in the same proportions. This prediction is known as *portfolio separation*. Let us determine whether portfolio separation obtained in one of the

experiments, namely, 981116, where CAPM pricing emerged in five of six periods.

Figure 16 documents that the individual portfolio holdings are at odds with the CAPM. It displays the distribution of final allocations in each of the six periods. Plotted are the individual holdings of security A (plus signs), expressed per dollar invested in risky securities. For comparison, the dollar weight on A in the market portfolio is also depicted (circles).¹⁶ CAPM predicts that the weight of Security A in the portfolio of each investor should be equal to that of the market portfolio. In the experiments, however, the weight of Security A in individual portfolios was as low as 0.00 and as high as 1.00, and appears quite random. The prediction of equal weights across all investors and equal the market portfolio weight, is not upheld in the data. Significantly, there is no improvement as time passes.

The same picture emerges in all other experiments. The situation is paradoxical, because, faced with portfolio holdings that are so at odds with the theory, one wonders why the pricing predictions do come about, because the pricing predictions rely crucially on portfolio separation. In the experiments, subjects evidently do not demand mean-variance optimal portfolios of risky securities (in equilibrium, only the market portfolio is mean-variance optimal). That is, investors purchase mean-variance suboptimal portfolios. How, then, can the market portfolio, which is the aggregate of investors' holdings, become mean-variance optimal, as required for CAPM pricing? We leave the resolution of this paradox for future work.

9 Conclusion

The experiments have demonstrated that the basic principles of modern asset pricing theory are very much at work in tightly controlled financial markets. Risk aversion is the driving force behind the trading, and markets tend to equilibrate, to the point that risk premia are determined solely by covariance with aggregate risk. These results are rendered more significant because subjects lacked the information that was necessary to deliberately set prices according to the theories we tested.

The experimental financial markets that are reported on in this paper were thick. Bid-ask spreads were generally tiny and the limit-order book extensive. This aided subjects' attempts to re-allocate their portfolios, unlike in our earlier experiments with far fewer subjects. The convergence to the CAPM equilibrium was not only far faster, but was not halted. This confirms our conjecture that thin markets create trade execution uncertainty which hampers price discovery.

The thick-market experiments involved up to 63 subjects. It would be interesting to know whether we can change the structure of the markets such that we can generate equilibration with far less subjects. Since the equilibration

¹⁶The dollar weight of the market portfolio changes slightly across periods, because of changes in the prices. The market portfolio has the same composition throughout the experiment: 7 units of security A and 11 of security B. Not all observations can be discerned in the plot, because many subjects held the same final portfolios (in terms of dollar weights).

problems appear to be caused by subjects' hesitance to re-allocate their portfolios in the face of trade execution uncertainty, altered designs would have to explicitly accomodate portfolio trading.

Despite the success of asset pricing theory in explaining the dynamics of our experimental financial markets, there appears to be a substantial amount of instability that is not quite thoroughly studied in the theory. In particular, the 990407 and 991111 experiments demonstrated how easily the market can be kicked off an equilibrium. These movements seem to be driven by sudden shifts in the subjective beliefs away from the objectively stated probabilities. The presence of subjective beliefs in itself does not violate the basic predictions of the theory, however. Indeed, we do find suggestive evidence that markets remain in equilibrium when we adjust for shifts in beliefs.

The importance of deviations of subjective beliefs from objective probabilities in explaining the experimental results calls for the development of an empirical methodology that allows for such deviations. Extant methodology does not do so. One rough technique that did was proposed here. Another, more decisive technique is developed in Bossaerts [1998], but works only in an environment where payoff information changes over time (in the experiments reported in this paper, information is static).

The evidence in favor of the hypergeometric belief structure suggests that it may be worthwhile to organize future experiments around random draws with replacement instead of the usual draws without replacement. It might be easier to accomodate subjects' natural inclinations since we are not interested in understanding how well subjects perceive risk, but instead, how risk perceptions lead to equilibration at the market level.

It is fair to characterize our experimental financial markets as being most of the time in the price discovery state. This contrasts with the theoretical hypothesis that financial markets equilibrate instantaneously. It also calls for an analysis of the stability of the equilibria that are the focus of modern asset pricing theory.

References

- Anderson, C., S. Granat, C. R. Plott and K. Shimomura (1999), "Global Instability in General Equilibrium: Experiments in the Scarf Environment," Manuscript, Caltech.
- Arrow, K. and F. Hahn (1971): *General Competitive Analysis*, San Francisco: Holden-Day.
- Benink, H. and P. Bossaerts (1998): "An Exploration Of Neo-Austrian Theory Applied To Financial Markets," Caltech Working Paper.
- Bossaerts, P. (1998): "The Dynamics Of Limited-Liability Securities Prices In Fallible Markets," Caltech Working Paper.
- Bossaerts, P., D. Kleiman and C. Plott (1998): "Price Discovery In Financial Markets: The Case Of The CAPM," Caltech Working Paper.
- Camerer, C. and K. Weigelt (1993): "Convergence in Experimental Double Auctions for Stochastically Lived Assets," in: *The Double Auction Market*, D. Friedman and J. Rust, eds., Reading, MA: Addison-Wesley.
- Clower, R. (1967): "A Re-Consideration of the Microfoundations of Monetary Theory," *Western Economic Journal* 6, 1-9.
- Debreu, G. (1959): *Theory Of Value*, New York: Wiley.
- Fama, E. and K. French (1992): "The Cross-section of Expected Stock Returns," *Journal of Finance* 47, 427-465.
- Forsythe, R., T. Palfrey and C. Plott (1984): "Futures Markets and Informational Efficiency: A Laboratory Examination," *Journal of Finance* 39, 955-81.
- Gallant, R.A. and G. Tauchen (1996), "Which Moments to Match?" *Econometric Theory* 12:657-681.
- Gouriéroux, C., A. Monfort and E. Renault (1993), "Indirect Inference," *Journal of Applied Econometrics* 8, S85-S118.
- Hansen, L.P. and K.J. Singleton (1982): "Generalized Instrumental Variables Estimation Of Nonlinear Rational Expectations Models," *Econometrica* 50: 1269-86.
- Harrison, M. and D. Kreps (1979): "Martingales and Arbitrage in Multiperiod Securities Markets," *Journal of Economic Theory* 20, 381-408.

- Kroll, Y., H. Levy and A. Rapaport (1988): "Experimental Tests of the Separation Theorem and the Capital Asset Pricing Model," *American Economic Review* 78, 500-519.
- Lei, V., C. Noussair, C. Plott (1998), "Non-Speculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality," Purdue University, Krannert Graduate School of Management Working Paper 1120.
- Levy, Haim (1997): "Risk and Return: An Experimental Analysis," *International Economic Review* 38, 119-149.
- Lintner, J. (1965): "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics* 47, 13-37.
- Lucas, R. (1982): "Interest Rates And Currency Prices In A Two-Country World," *Journal of Monetary Economics* 10, 335-59.
- Mehra, R. and E. Prescott (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15, 145-61.
- Mehra, R. and R. Sah (1998): "Can Small Fluctuations In Investors' Subjective Preferences Induce Large Volatility In Equity Prices?" UCSB Working Paper.
- Mossin, J. (1966): "Equilibrium In A Capital Asset Market," *Econometrica* 34, 768-83.
- Nielsen, L. (1988): "Uniqueness of Equilibrium in the Classical Capital Asset Pricing Model," *Journal of Financial and Quantitative Analysis* 23, 329-336.
- Noussair, C., C. Plott and A. Rezin (1995), "An Experimental Investigation of the Patterns of International Trade," *American Economic Review* 85, 462-491.
- Plott, C.R. and S. Sunder (1988): "Rational Expectations and the Aggregation of Diverse Information in Laboratory Security Markets," *Econometrica* 56, 1085-1118.
- Roll, R. (1977): "A Critique Of The Asset Pricing Theory's Tests, Part I: On The Past And Potential Testability Of The Theory," *Journal Of Financial Economics* 4, 129-76.
- Rubinstein, M. (1994): "Implied Binomial Trees," *Journal Of Finance* 49, 771-818.
- Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk," *Journal of Finance* 19, 425-442.
- Smith, V., G. Suchanek and A. Williams (1988): "Bubbles, Crashes and Endogenous Expectations in Experimental Spot Markets," *Econometrica* 56, 1119-51.

Table 1: Experimental Design Data

Experiment	Subject Category (Number)	Signup	Endowments			Cash	Loan	Exchange Rate \$/franc
		Reward (franc)	A	B	Notes	(franc)	Repayment (franc)	
981007	30	0	4	4	0	400	1900	0.03
981116	23	0	5	4	0	400	2000	0.03
	21	0	2	7	0	400	2000	0.03
990211	8	0	5	4	0	400	2000	0.03
	11	0	2	7	0	400	2000	0.03
990407	22	175	9	1	0	400	2500	0.03
	22	175	1	9	0	400	2400	0.04
991110	33	175	5	4	0	400	2200	0.04
	30	175	2	8	0	400	2310	0.04
991111	22	175	5	4	0	400	2200	0.04
	23	175	2	8	0	400	2310	0.04

Table 2: Test Of Stochastic Attraction Towards CAPM Equilibrium Against Random Walk Pricing

Experiment	Attraction Coefficient κ		
	Estimate ^a	Critical Value ^b	
		5%	10%
981007	-0.103**	-0.013	-0.009
981116	-0.081**	-0.009	-0.004
990211	-0.462**	-0.219	-0.081
990407	-0.016**	-0.009	-0.006
991110	-0.013	-0.119	-0.033
991111	-0.056**	-0.014	-0.007

^aMeaning of superscripts: ** = significant at the 5% level.

^bBased on 200 bootstrapped samples of the same size as used to estimate κ .

Table 3: Test Of Stochastic Attraction Towards Arrow Debreu Equilibrium Against Random Walk Pricing

Experiment	Mean Transition Probability π			
	Estimate ^a	Critical Value ^b		
		5%	90%	95%
981007	0.88**	0.31	0.61	0.59
981116	0.87**	0.25	0.48	0.51
990211	0.84**	0.39	0.79	0.83
990407	0.89**	0.64	0.83	0.85
991110	0.88*	0.64	0.88	0.89
991111	0.87**	0.36	0.70	0.73

^aMeaning of superscripts: ** = significant at the 95% level; * = significant at the 90% level.

^bBased on 200 bootstrapped samples of the same size as used to estimate π .

Table 4: Market Thickness And Bid-Ask Spreads

Experiment ^a	Subject Number	Markets								
		A			B			Notes		
		\bar{x}^b	m^c	$\hat{\sigma}^d$	\bar{x}	m	$\hat{\sigma}$	\bar{x}	m	$\hat{\sigma}$
981007	30	9.5	5	10.1	4.4	3	5.9	2.7	1	2.6
981116	44	5.6	3	6.6	4.6	2	5.7	2.3	1	2.4
990211	19	10.9	7	10.3	8.8	5	9.4	5.0	4	3.0
990407	44	4.8	3	5.7	5.0	3	5.3	1.6	1	1.1
991110	63	5.1	3	6.1	8.3	4	9.8	3.4	1	4.1
991111	45	5.8	4	5.7	3.9	3	3.8	2.0	1	1.4

^aExcludes first period in each experiment.

^bSample average (in francs).

^cSample median (in francs).

^dSample standard deviation (in francs).

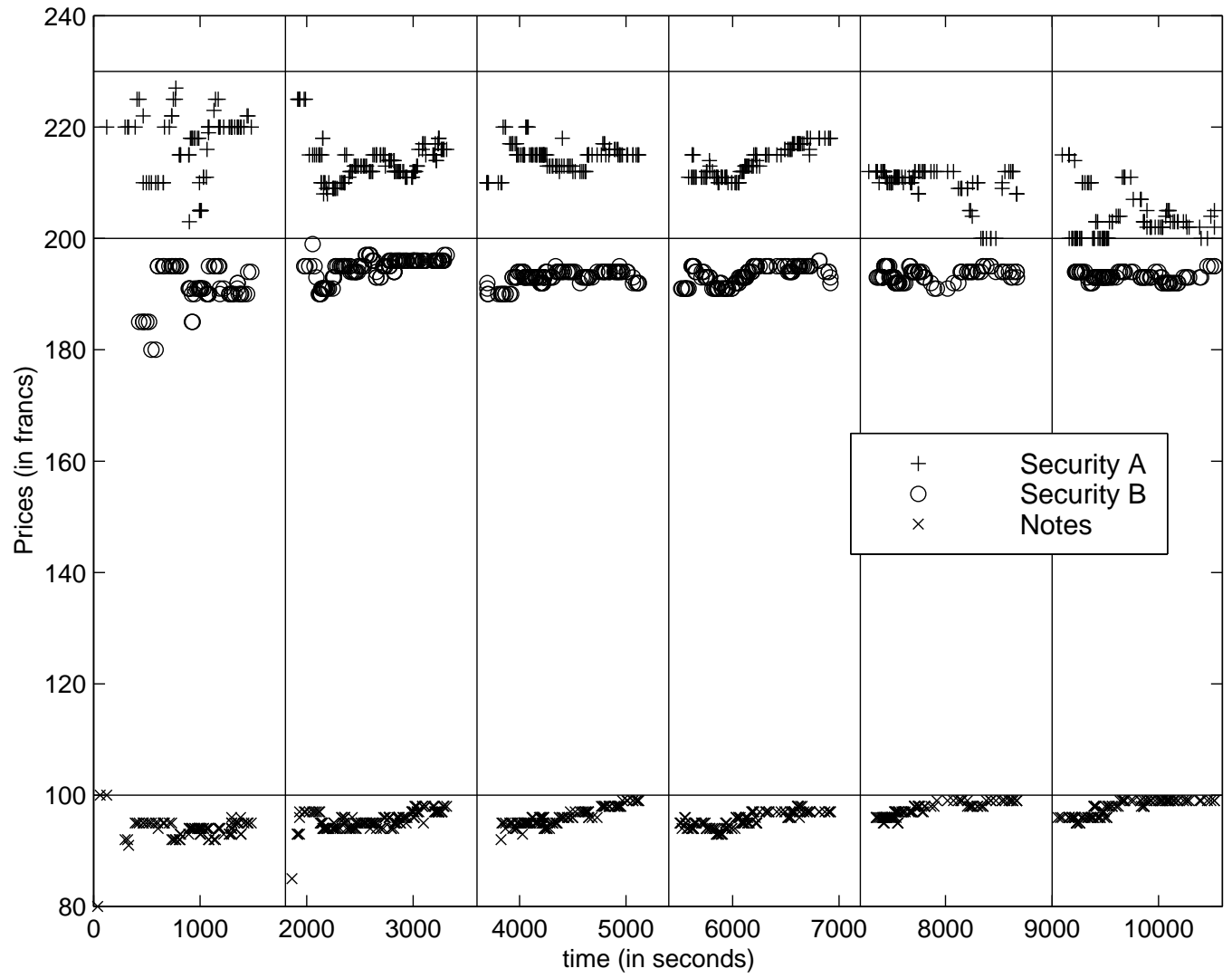


Figure 1: Evolution of securities prices in the 981007 experiment.

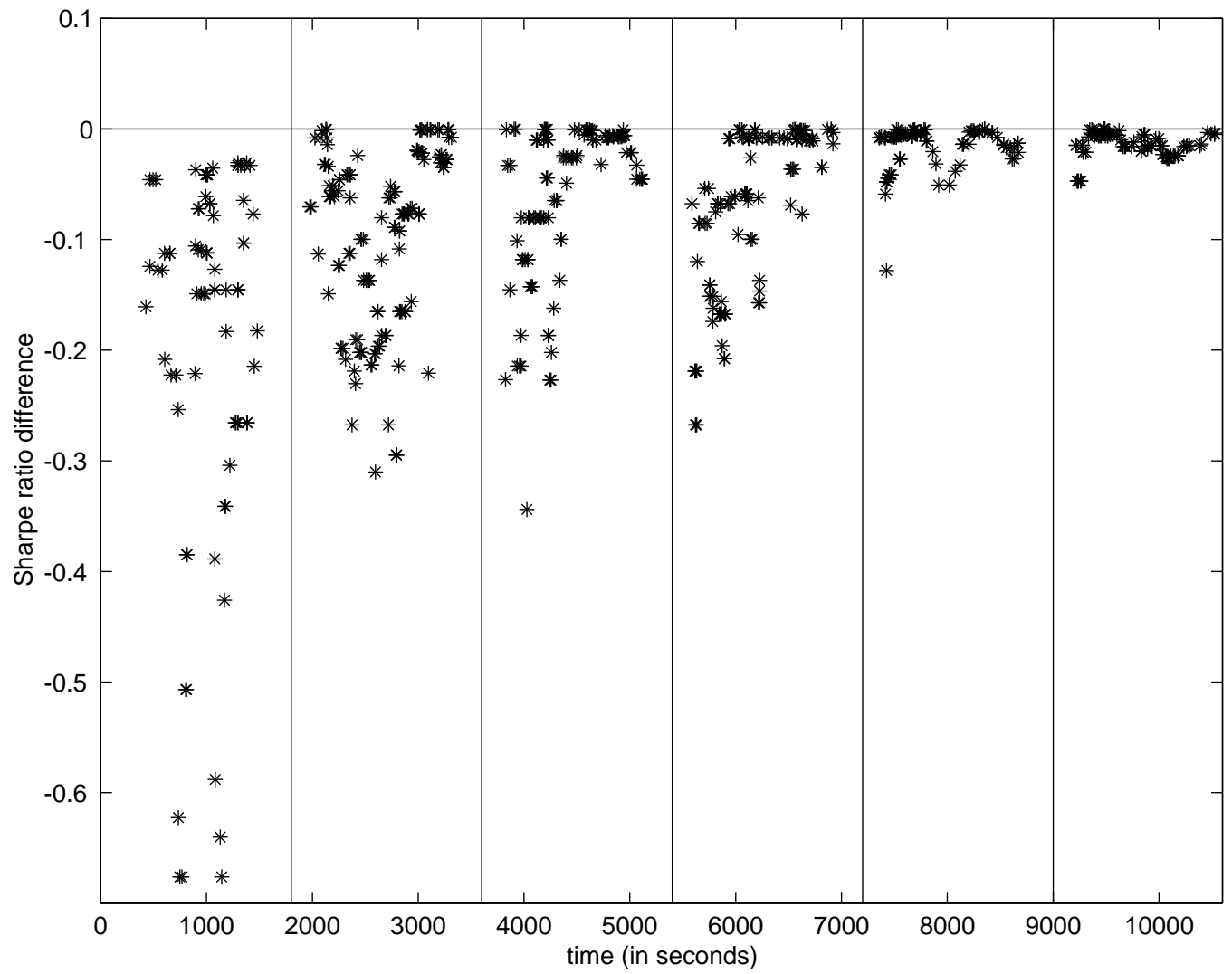


Figure 2: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 981007 experiment.

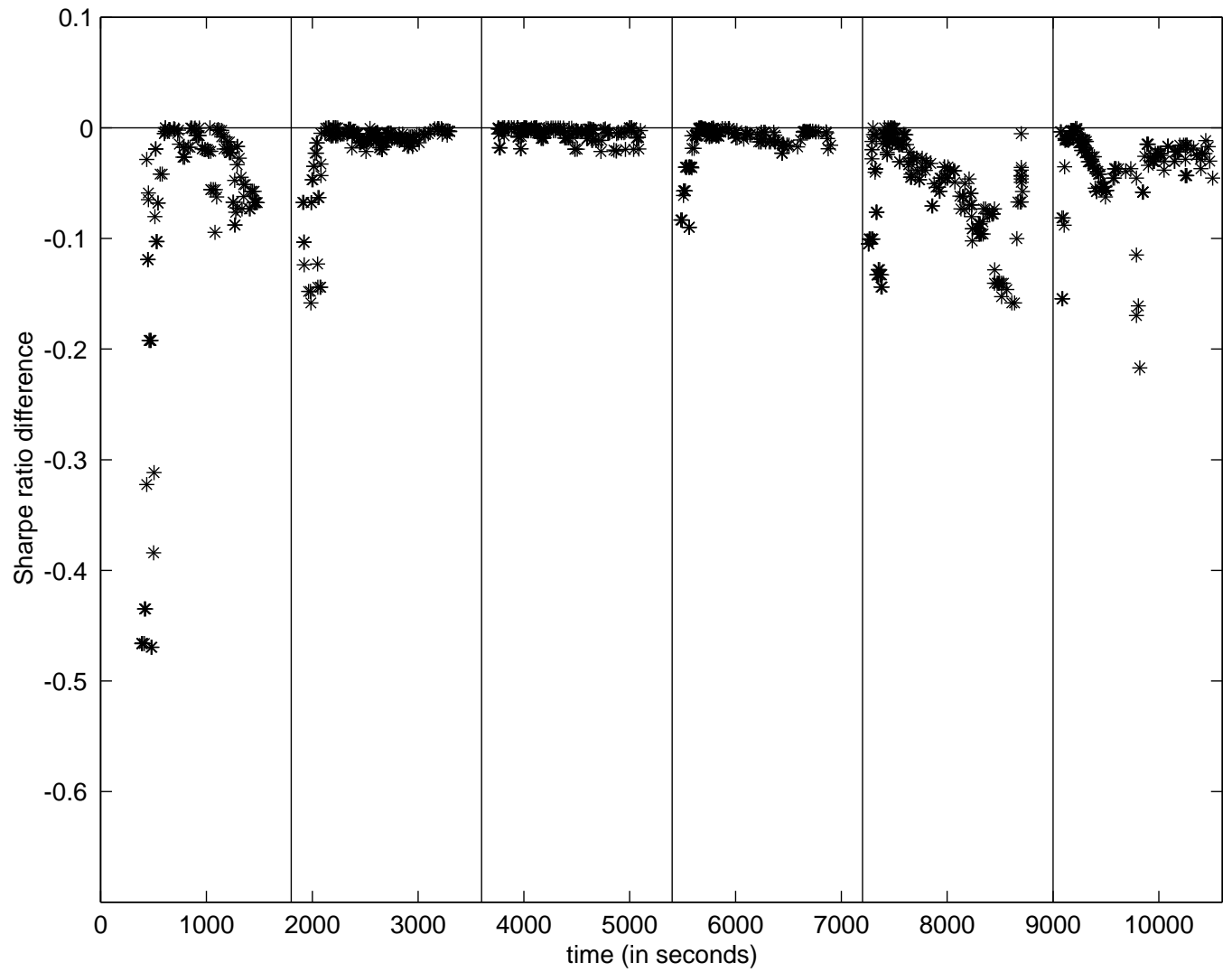


Figure 3: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 981116 experiment.

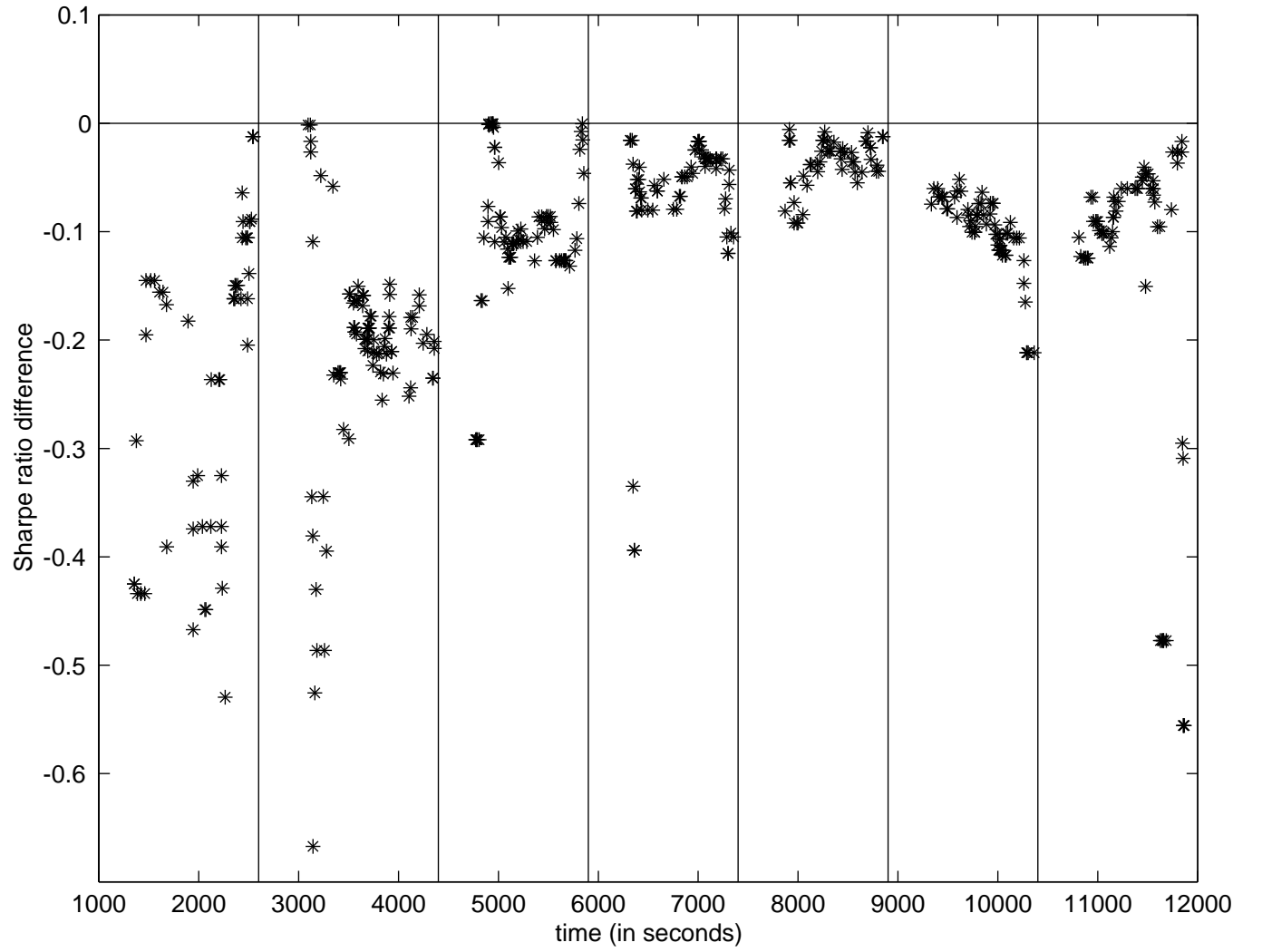


Figure 4: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 990211 experiment.

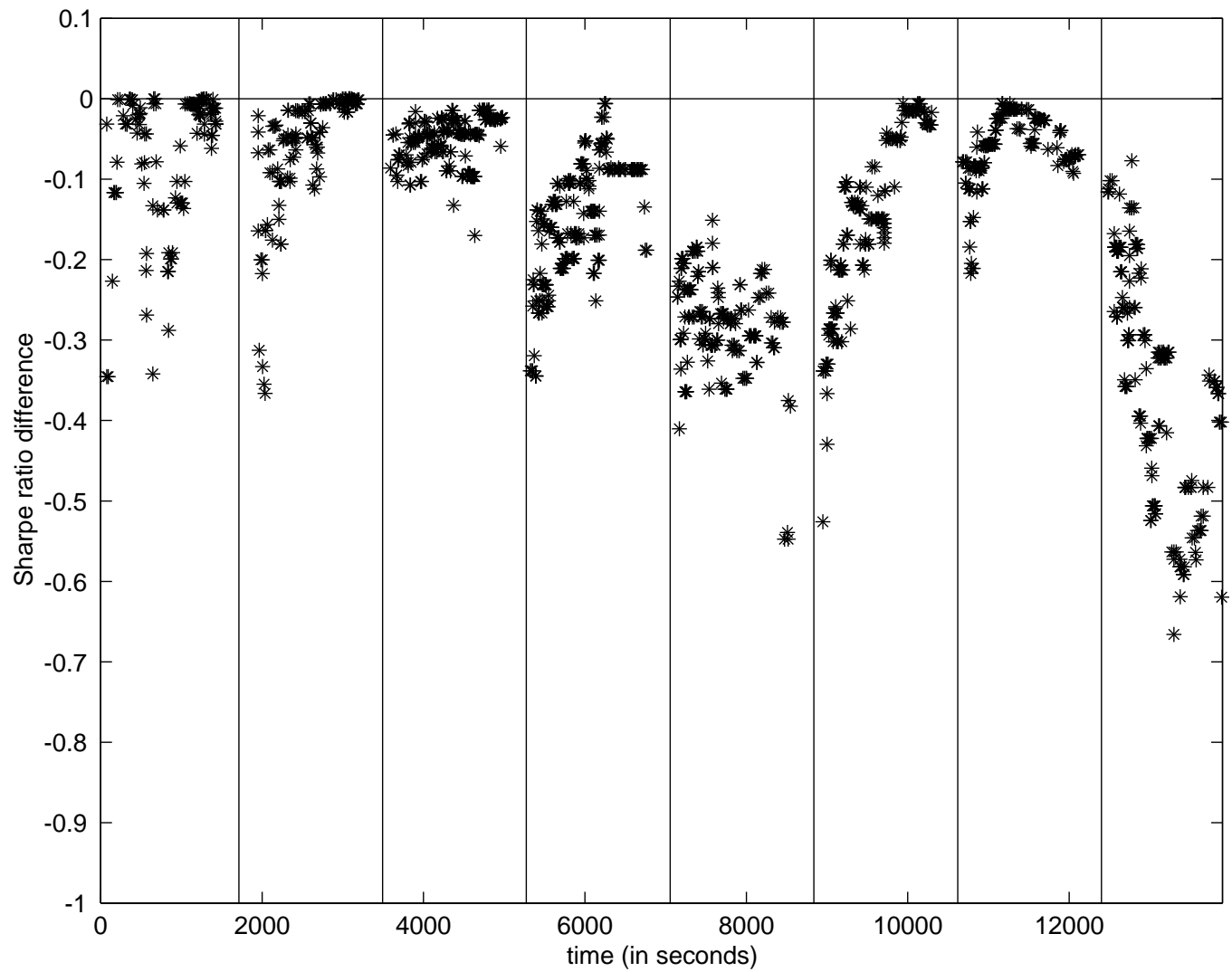


Figure 5: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 990407 experiment.

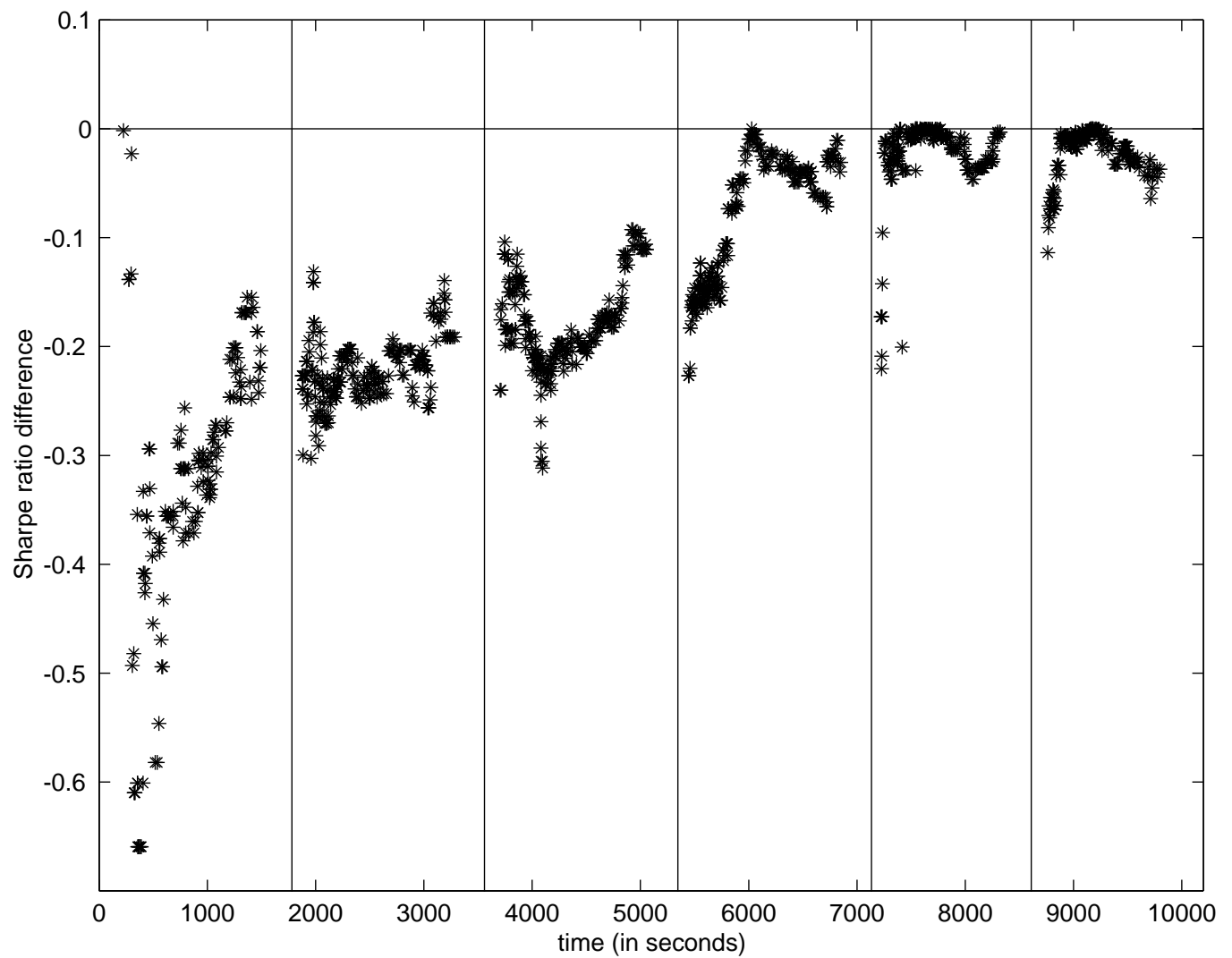


Figure 6: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 991110 experiment.

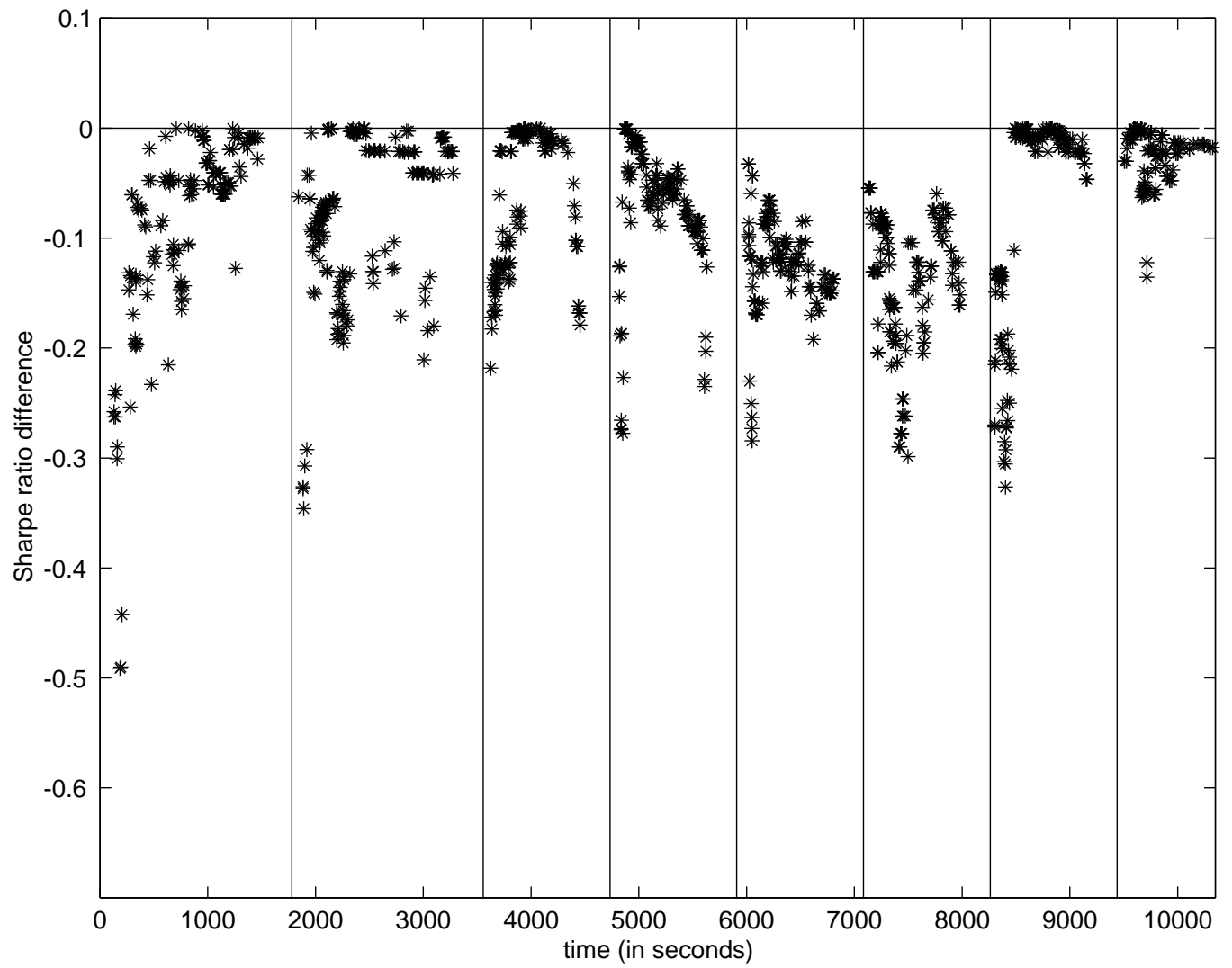


Figure 7: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 991111 experiment.

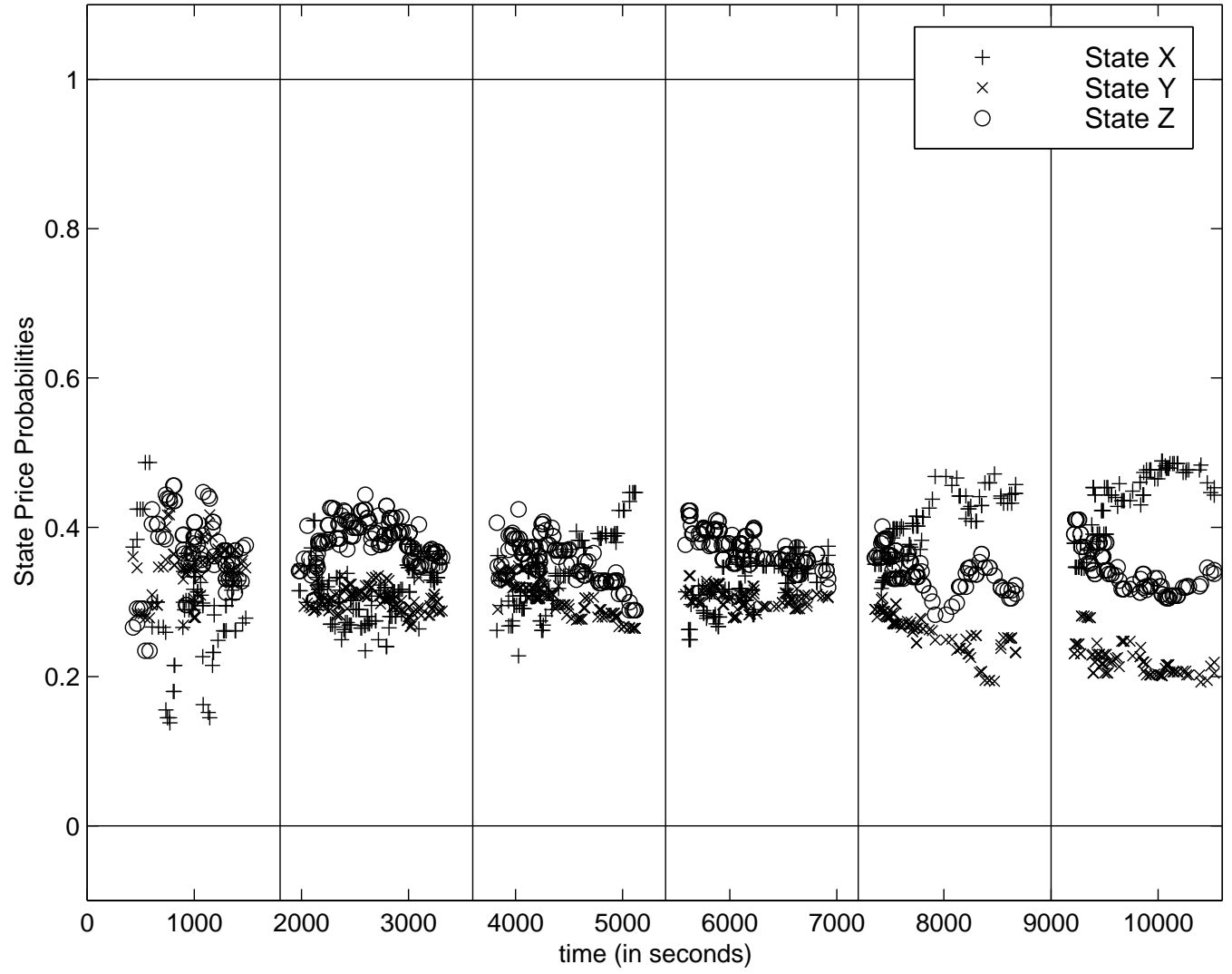


Figure 8: Evolution of the state-price probabilities in the 981007 experiment.

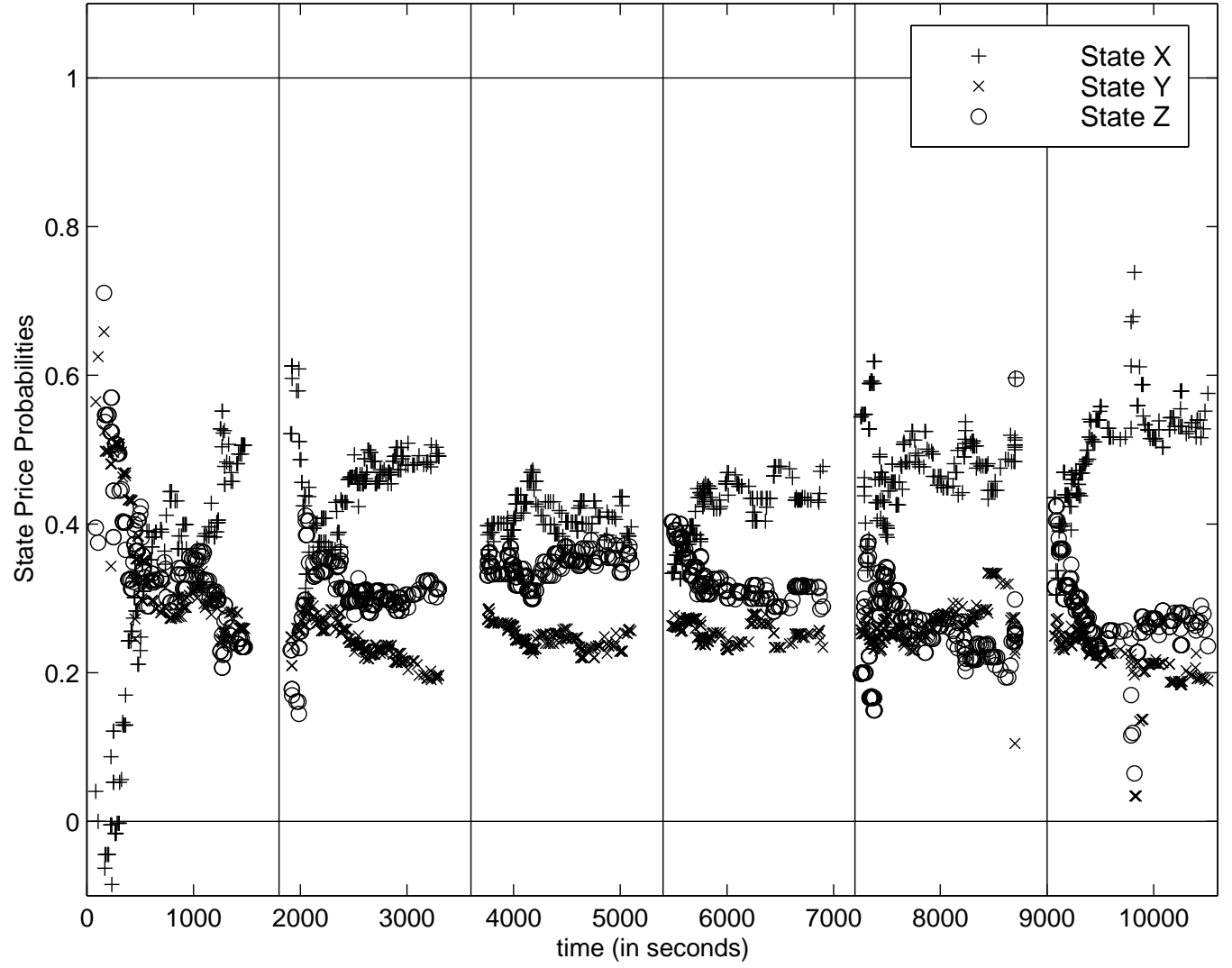


Figure 9: Evolution of the state-price probabilities in the 981116 experiment.

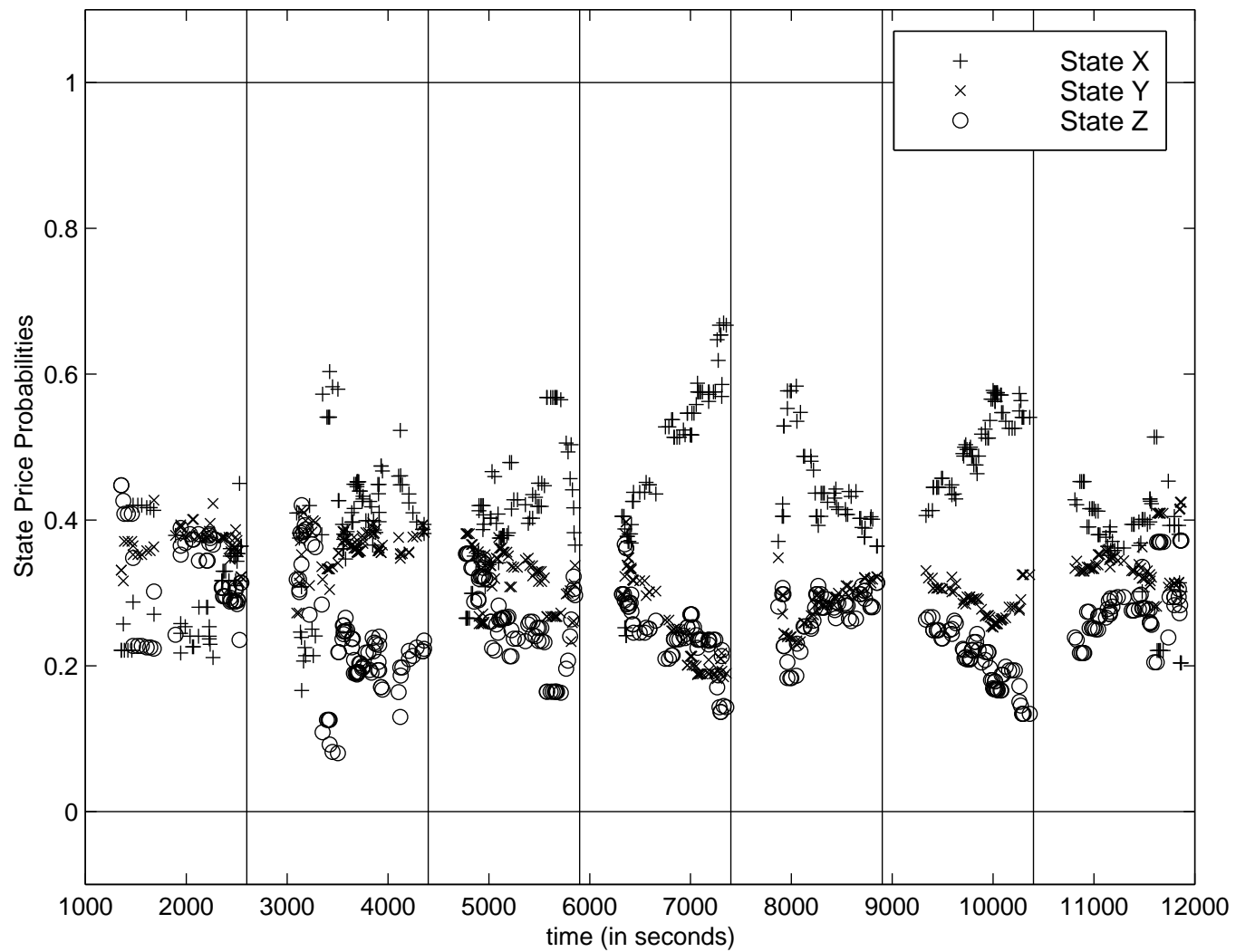


Figure 10: Evolution of the state-price probabilities in the 990211 experiment.

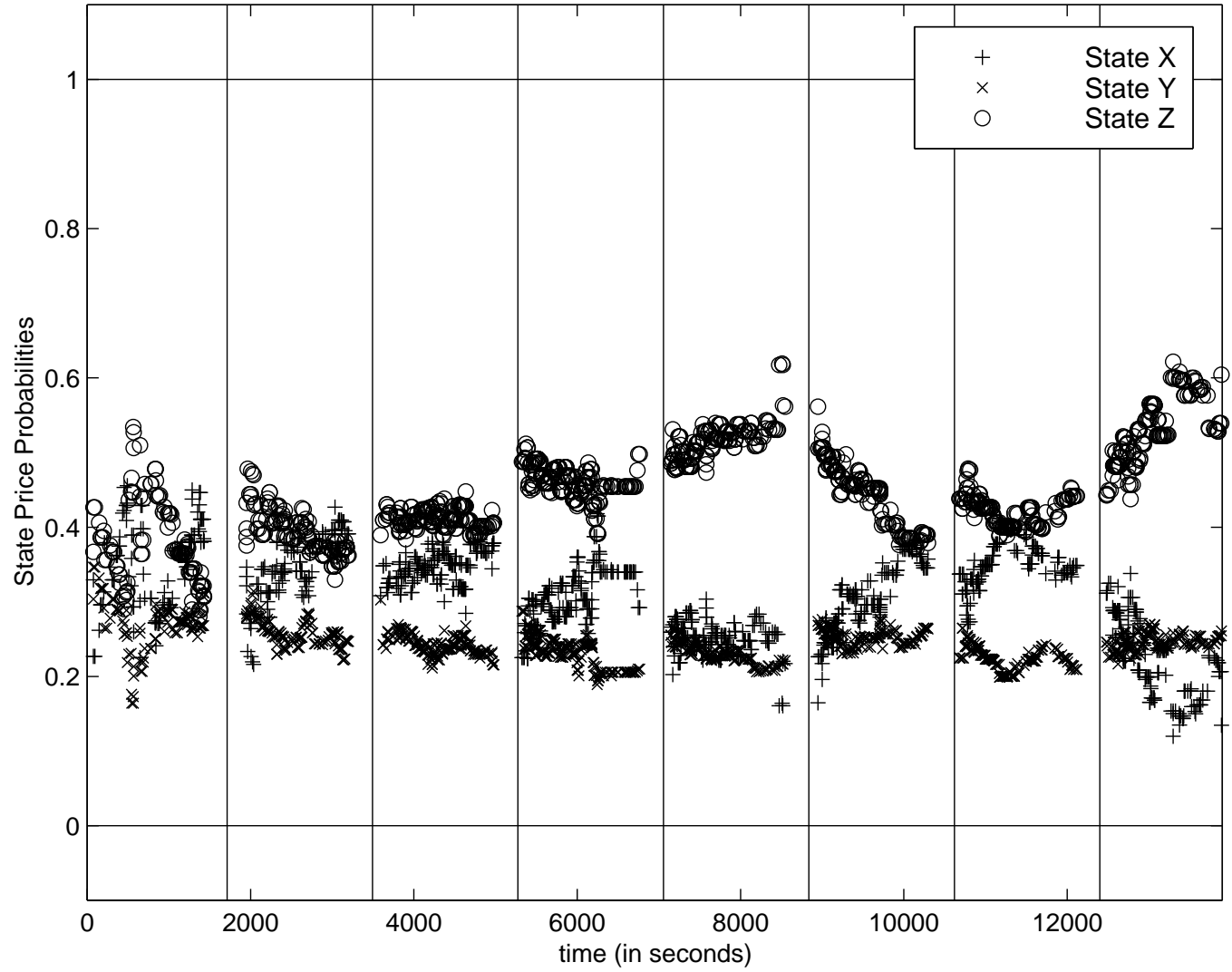


Figure 11: Evolution of the state-price probabilities in the 990407 experiment.

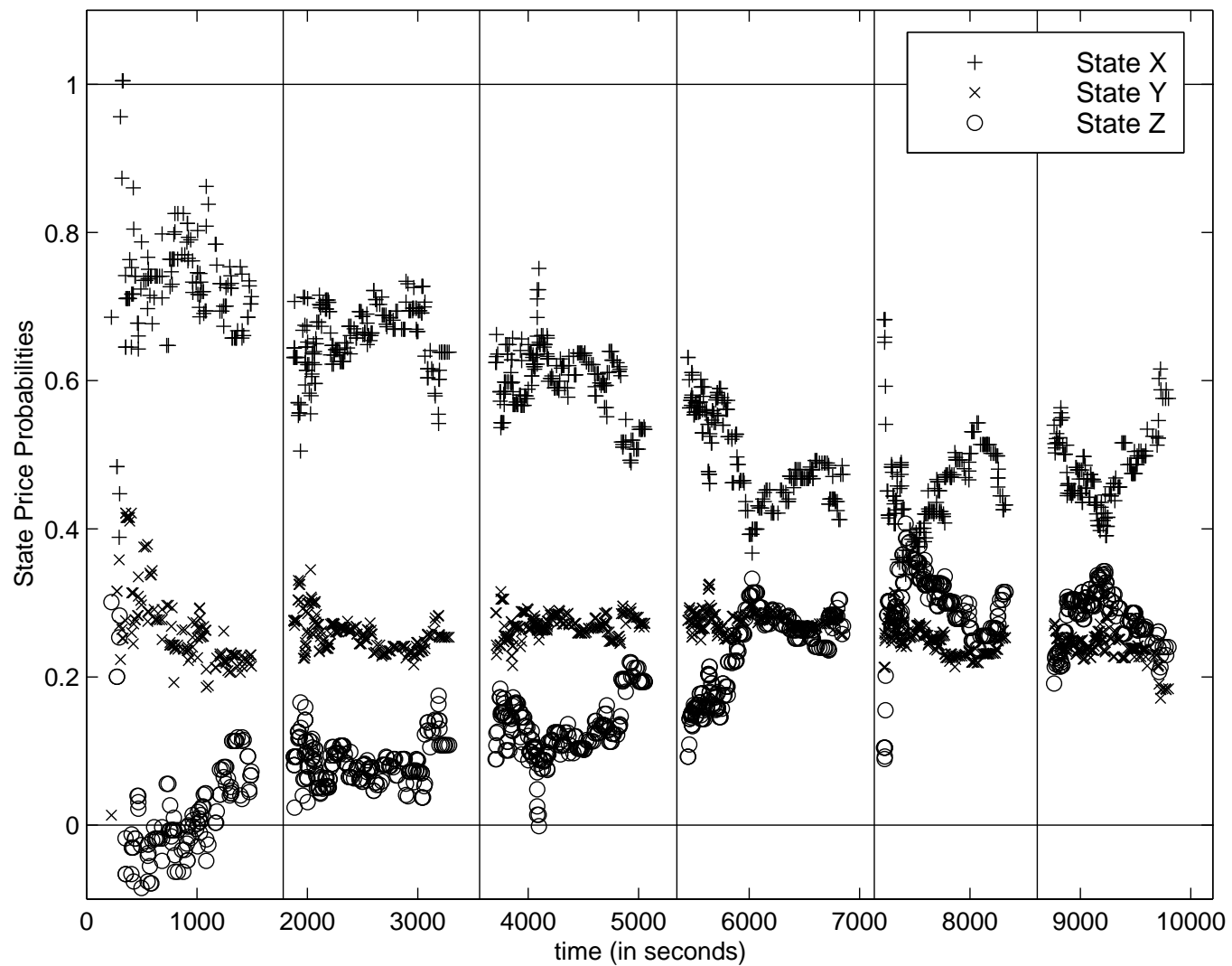


Figure 12: Evolution of the state-price probabilities in the 991110 experiment.

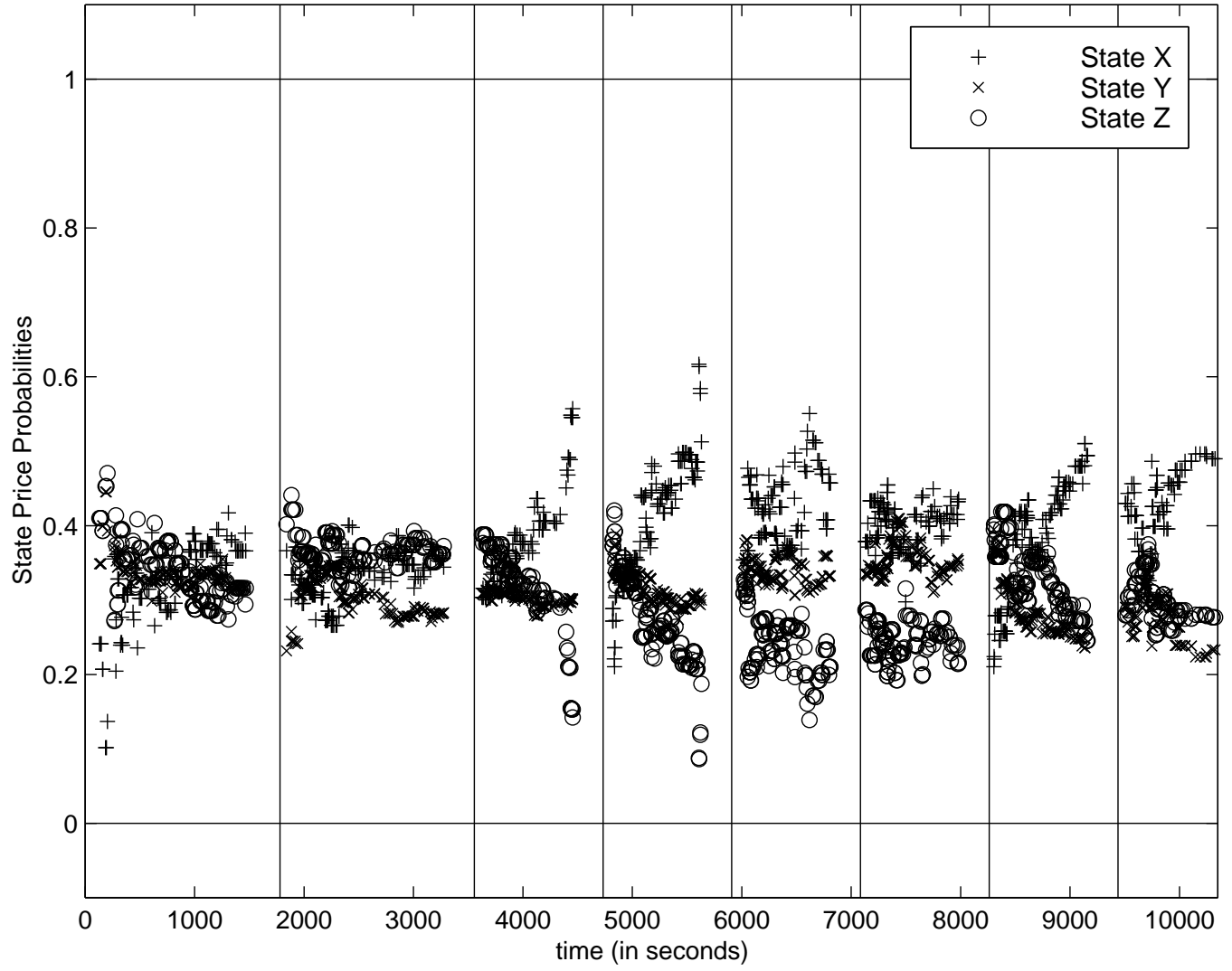


Figure 13: Evolution of the state-price probabilities in the 991111 experiment.

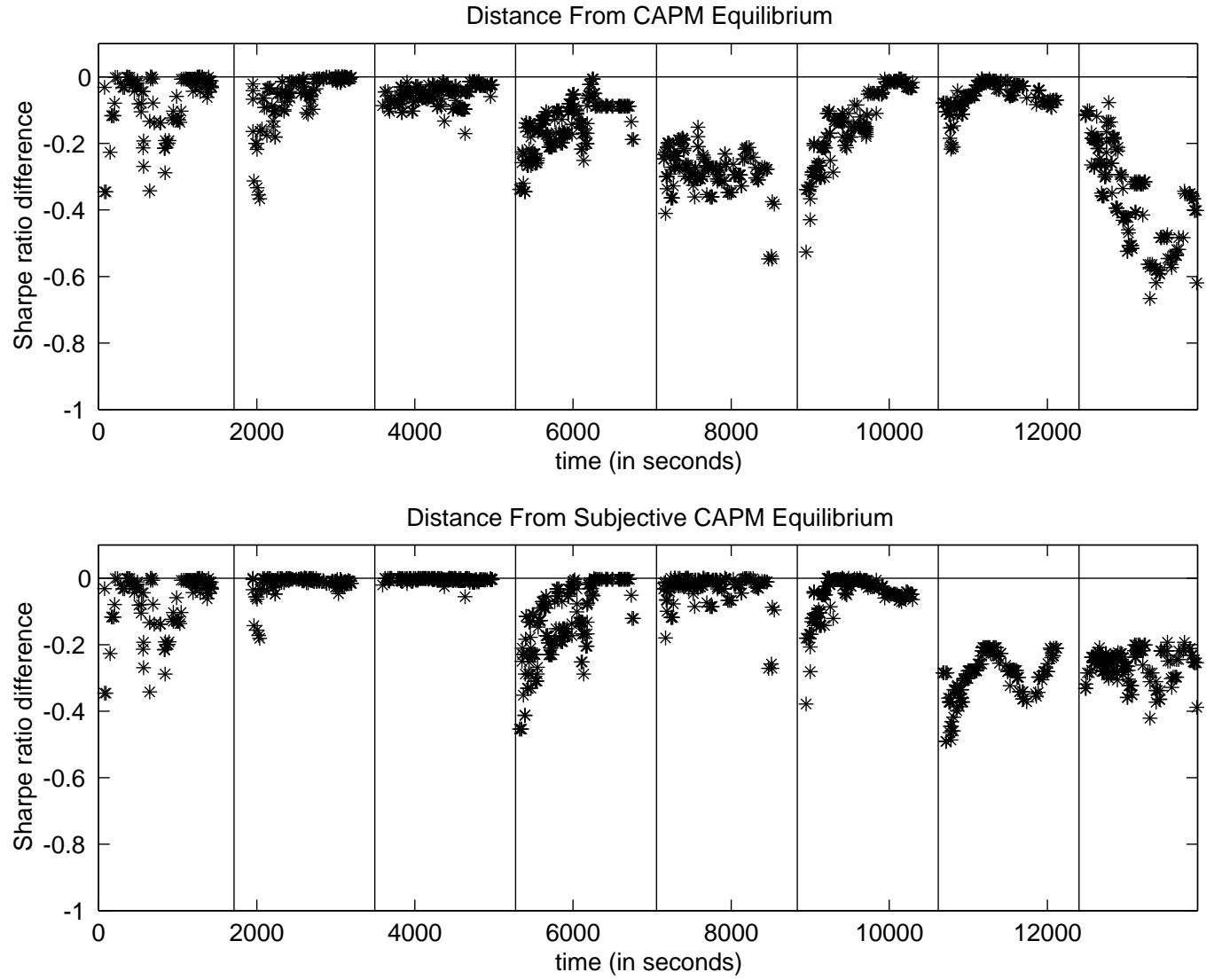


Figure 14: Top Panel: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 990407 experiment. Bottom Panel: Same evolution, based on the belief that states were drawn without replacement from an urn with 12 balls, 4 of each state.

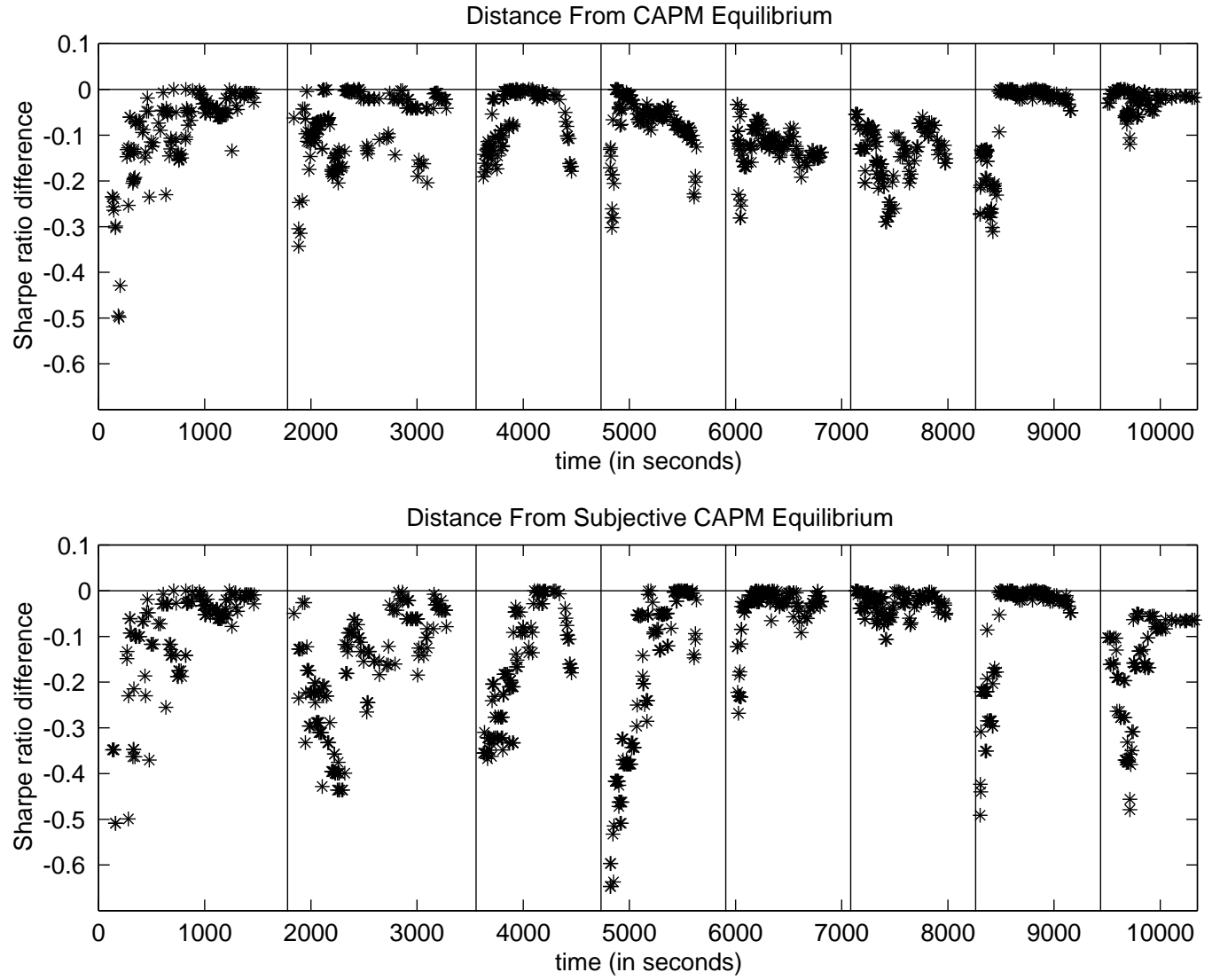


Figure 15: Top Panel: Evolution of the difference between the Sharpe ratio (expected return divided by standard deviation) of the market portfolio and the maximum possible Sharpe ratio, 991111 experiment. Bottom Panel: Same evolution, based on the belief that states were drawn without replacement from an urn with 12 balls, 4 of each state.

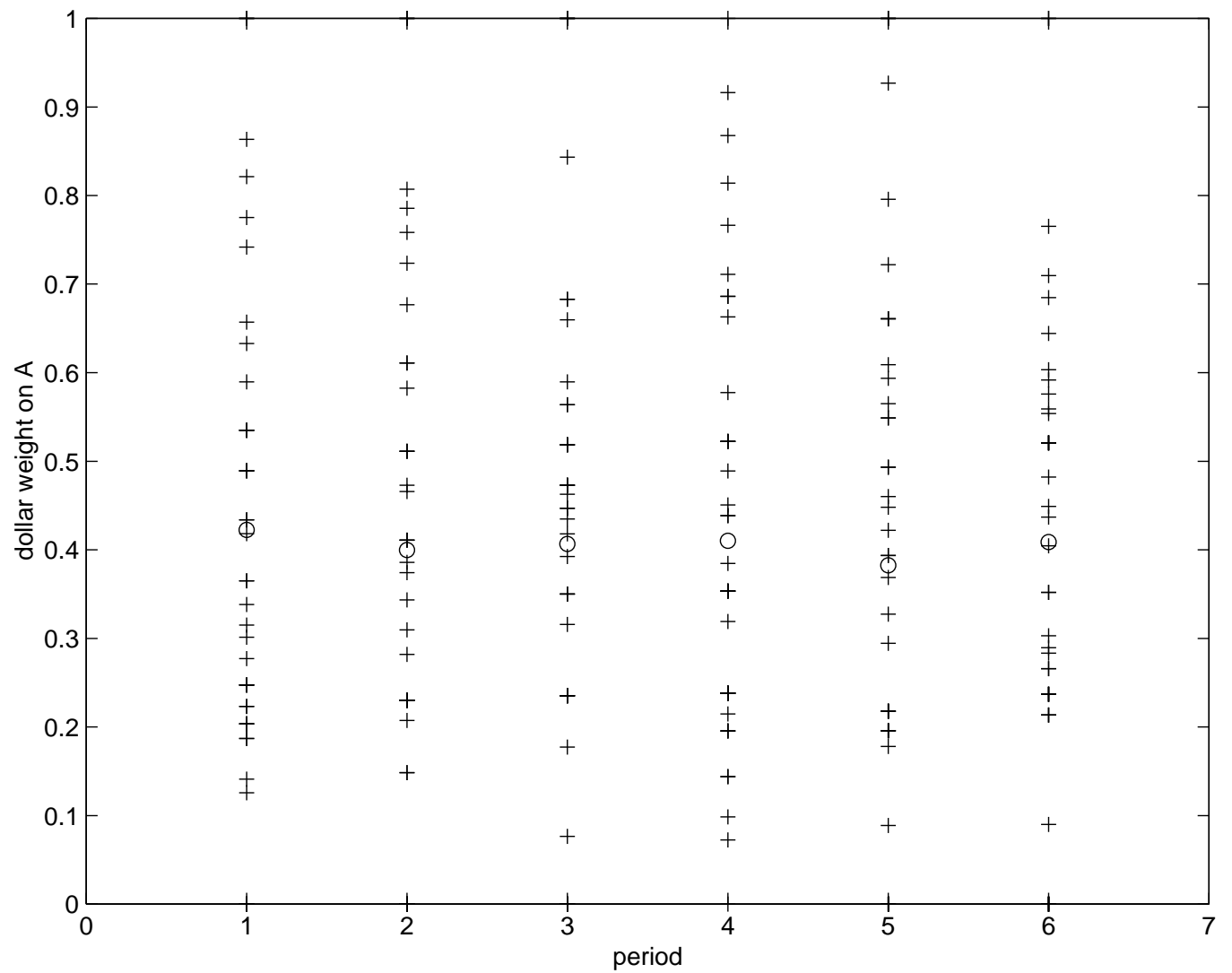


Figure 16: Distribution of dollar weights (plus signs) in individual portfolios of risky securities at the end of each period, experiment 981116. For comparison, market portfolio weights are also depicted (circles). According to CAPM, all individual portfolio weights should be equal to market portfolio weights.